IMPLEMENTING SUPER-RESOLUTION PALM MICROSCOPY IN FISSION YEAST

By

Helen Elizabeth Harcourt Armes

SUBMITTED FOR THE
DEGREE OF DOCTOR OF
PHILOSOPHY

The University of Sussex

September 2016

Declaration

I hereby declare that this thesis has not been and will not be, submitted in whole or in part to another University for the award of any other degree.

Signature:.....

Date:.....

This thesis is dedicated to my family and friends

Acknowledgements

This has been a very cross-disciplinary project with contributions from a number of different labs and I have had support from a lot of different people.

I would like to start by thanking my supervisor Professor Tony Carr for having me in his group and suggesting these projects. He has been a very supportive, always showing interest in what I have done and making suggestions for next steps.

I would like to thank Professor David Klenerman (University of Cambridge) for letting me spend a year in his lab learning the ropes of super-resolution microscopy. I would also like to thank Dr Steve Lee and Dr Matthieu Palayret for teaching me to use the microscope and Dr David Lando (Prof. Ernest Laue's Group, University of Cambridge) for assisting with learning to handle *S. pombe*. I'd also like to acknowledge all the members of the Klenerman lab at that time who made me feel very welcome and the other members of the Laue lab who were involved in the PALM collaboration.

At the University of Sussex I would like to thank Dr Adam Watson who has given a lot of his time to help me with biological work and to provide me with the strains I needed. I would like to thank Dr Manisankar Maiti for tireless work on aligning, calibrating and maintaining the instruments I have used, I never would have got the FCS data without him. Also Dr Mark Osborne for his input into setting up the FCS system and interpreting my results. I would like to thank Dr Alex Herbert, the writer of PeakFit and diverse other plugins I have used for my data analysis, not only did he devote a lot of time to developing these tools but also a lot of time to explaining them to me. I would like to thank Dr Thomas Etheridge who always had at least ten suggestions of different things to try whenever I got stuck.

I would like to thank my co-supervisor Dr Eva Hoffman for patiently and persistently

encouraging me to start writing, and for reading and correcting first drafts. I would also like to thank Dr Jo Murray for offering extra support in proof reading this thesis

I would like to thank Dr Dorota Dziadkowiec and Anna Barg for sharing strains, data and encouragement with me for the Rrp1/2 collaboration.

I would like to thank all other members of both the Carr and Murray lab groups for being generally friendly, knowledgeable and always ready to provide advice.

I would also like to thank the friends and family who have kept me sane for the last four years. In particular my partner Simon who has provided much support and many cups of tea and my college wives who have been there for me almost constantly via Facebook, encouraging me when I felt like giving up.

UNIVERSITY OF SUSSEX

Helen Elizabeth Harcourt Armes
A thesis submitted for the degree of Doctor of Philosophy

IMPLEMENTING SUPER-RESOLUTION PALM MICROSCOPY IN FISSION YEAST

SUMMARY

Fluorescence microscopy is a popular biological technique because it allows the study of cells in great detail. However, the resolution achievable is limited by the diffraction properties of light, meaning that fine detail cannot be resolved. Various super-resolution microscopy methods have been developed to break this resolution limit. This thesis focuses on the single molecule localisation microscopy techniques. My host laboratory focuses on DNA replication and repair pathways using the model organism *Schizosaccharomyces pombe* (fission yeast). The aim of this thesis is thus to apply the technique of photo-activatable localisation microscopy (PALM) to specific biological questions in order to establish its benefits and limitations.

In theory, in PALM every molecule will be imaged once and, as such, could be counted. So far this has been largely limited to membrane proteins. Using a combination of artificially created fluorescent oligomers, endogenous ribonucleotide reductase proteins tagged with mEos and computer simulations I studied the feasibility of counting highly expressed cytoplasmic proteins and assigning them to complexes of known or unknown stoichiometry. I established that density of expression is a significant limiting factor when using PALM to resolve complex stoichiometry.

I thus went on to develop a variation of fluorescence correlation spectrometry to study the same protein complexes to see if we could determine their stoichiometry by diffusion speed. I established that the technique could differentiate between quite small changes in size. However the endogenous complex did not respond well to the fluorophore used so I was not able to establish its size.

Using the PALM system I also studied a biological molecule, Rrp2, which was expressed at such low levels it was not possible to observe with conventional fluorescence microscopy. I established that we were able to observe this protein at endogenous levels and characterised its behaviour in response to stress.

Table of Contents

Table of Figures

List of Tables

List of Abbreviations

Abbreviation	Meaning
2D	Two-dimensional
3D	Three-dimensional.
5-FOA	5-Fluoroorotic Acid
Amp	Ampicillin sodium salt
APD	Avalanche photo-diode
BODIPY	boron-dipyrromethene
CCD	Charge-coupled device
CMOS	complementary metal-oxide-semiconductor
dATP	Deoxy-adenosine triphosphate
dCTP	Deoxy-cytidine triphosphate
dGTP	Deoxy-guanosine triphosphate
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNDP	Deoxy-ribonucleotide diphosphate
dNMP	Deoxy-ribonucleotide monophosphate
dNTP	Deoxy-ribonucleotide triphosphate
DSB	Double strand break
dTTP	Deoxy-thymidine triphosphate
EDTA	Ethylenediaminetetraacetic acid
EMCCD	Electron multiplying charge-coupled device
EMM	Edinburgh minimal media
FCS	Fluorescence correlation spectroscopy
FOV	Field of view
FP	Fluorescent protein
FPALM	Fluorescence photo-activatable localisation microscopy
FRET	Förster resonance energy transfer
FWHM	Full width at half maximum
G418	Geneticin disulphate
GME	Gaussian mask estimator
HR	Homologous recombination
HU	Hydroxyurea
Hygro	Hygromycin
IR	Ionising Radiation
LB	Lysogeny broth
LBA	Lysogeny broth agar
LSE	Least squares estimation
MLE	Maximum likelihood estimation
MLEwT	Maximum likelihood estimation with a theoretical PSF
MMS	Methyl methanesulfonate
NA	Numerical aperture
Nat	Nourseothricin sulphate
NDP	Ribonucleotide diphosphate
NHEJ	Non-homologous end joining
PA	Photo-activatable

PAFP	Photo-activatable fluorescent proteins
PALM	Photo-activation localisation microscopy
PBS	Phosphate buffered saline
PBST	Phosphate buffered saline with tween
PC PALM	Pairwise correlation PALM
PEG	Polyethylene glycol
PS	Photo-switchable
PSF	Point spread function
QD	Quantum dots
RESOLFT	Reversible saturable optical linear transitions
RFB	Replication fork barrier
RPA	Replication protein A
SMLM	Single molecule localisation microscopy
SNR	Signal to noise ratio
SPAD	Single photon avalanche photodiode
ssDNA	Single stranded DNA
SSIM	Saturated structured-illumination microscopy
STED	Stimulated emission depletion
STORM	Stochastic optical reconstruction micrroscopy
SUMO	Small ubiquitin-like modifier
TE	Tris-EDTA buffer
TMR	tetramethylrhodamine
UV	Ultra Violet
WLE	Weighted least-squares
YE	Yeast extract
YEA	Yeast extract agar

1. Introduction

1.1 General introduction and thesis aims

Over the last few decades the biological sciences have seen many advances in instrumentation, technology and computing-driven solutions that allow observation of different aspects of cellular and molecular processes in ever greater detail. In particular, the developments surrounding fluorescence microscopy have been revolutionary in improving our ability to visualise the processes occurring inside cells. More recently the breaking of the diffraction limit in microscopy by so-called super-resolution technologies has opened up a whole new realm of possibilities. However all technologies have limitations as well as benefits and it is important to know which tools are suitable for answering which questions. The main aim of this thesis has been to use biological problems to investigate the limitations and benefits of the super-resolution technique photo-activated localisation microscopy (PALM). PALM has been used quantitatively by a number of research groups but mainly for low abundance, membrane localised proteins. One of the aims of this thesis has been to see whether the stoichiometry of abundant pan-cellular proteins can be addressed by PALM. To achieve this, the lab's experience in studying the ribonucleotide reductase (RNR) complex was exploited. A second, contrasting aim was to study a low abundance protein. For this I exploited the nuclear protein Rrp2, which is expressed at such low abundance that it is difficult to image it at endogenous levels using conventional fluorescence microscopy. The aim of this thesis was thus to explore using the improved resolution of PALM to visualise Rrp2 and investigate its interactions with DNA.

1.2 Super-resolution microscopy

1.2.1 Light microscopy

Microscopy has long been a mainstay of the biological sciences. Sometime in the 1590's Dutch spectacle makers Zacharias and Hans Jansen put together the first compound microscope (Croft, 2006). This microscope consisted of three draw tubes with a bi-convex eyepiece lens at one end and a plano-convex objective lens at the other, the microscope could be focused by sliding one of the draw tubes in or out achieving a magnification of up to ten times when fully extended. This magnification allowed visualisation of much smaller objects than had previously been achievable. In the 17th century the British born microscopist Robert Hook used compound microscopes of his own design to study a variety of objects from the surfaces of wood and leaves down to lice and fleas. He appears to have been the first to publish a depiction of a micro-organism, having observed a mould growing on the cover of a leather bound book (Gest, 2007). He also coined the term "cell" to describe the structures he observed in a piece of cork, which he compared to the sleeping rooms in a monastery. In 1665 he published a book called Micrographia which was filled with copperplate illustrations produced from his observations as well as a detailed description of the microscopes he built (Hooke, 1665). Around the same time as Hooke was working the Dutch microscopist Anton Van Leeuwenhoek started using single lens microscopes (which suffered less from aberrations caused by the lenses) and was the first to observe a number of different types of cell (Gest, 2007). Since then microscopy has been essential in enabling the study of single cells and direct observation of how they behave, and it has transformed the field of biology.

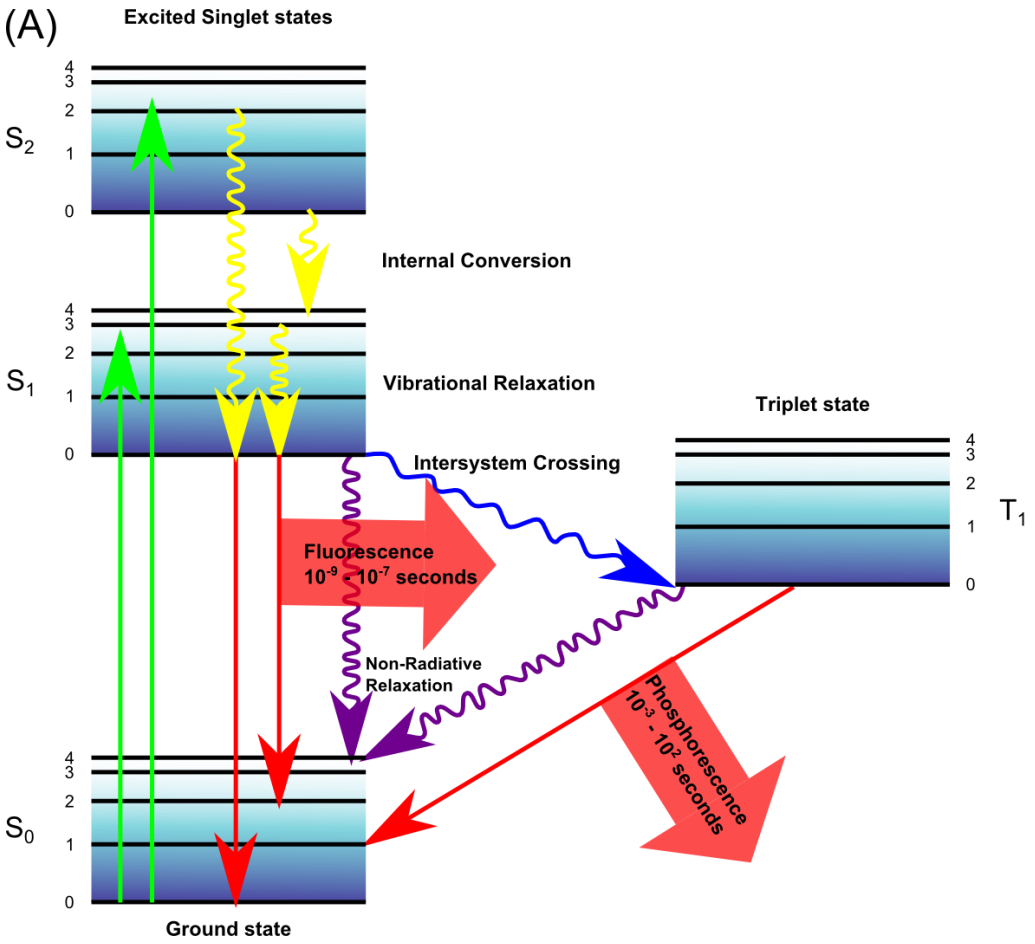
1.2.2 Fluorescence microscopy

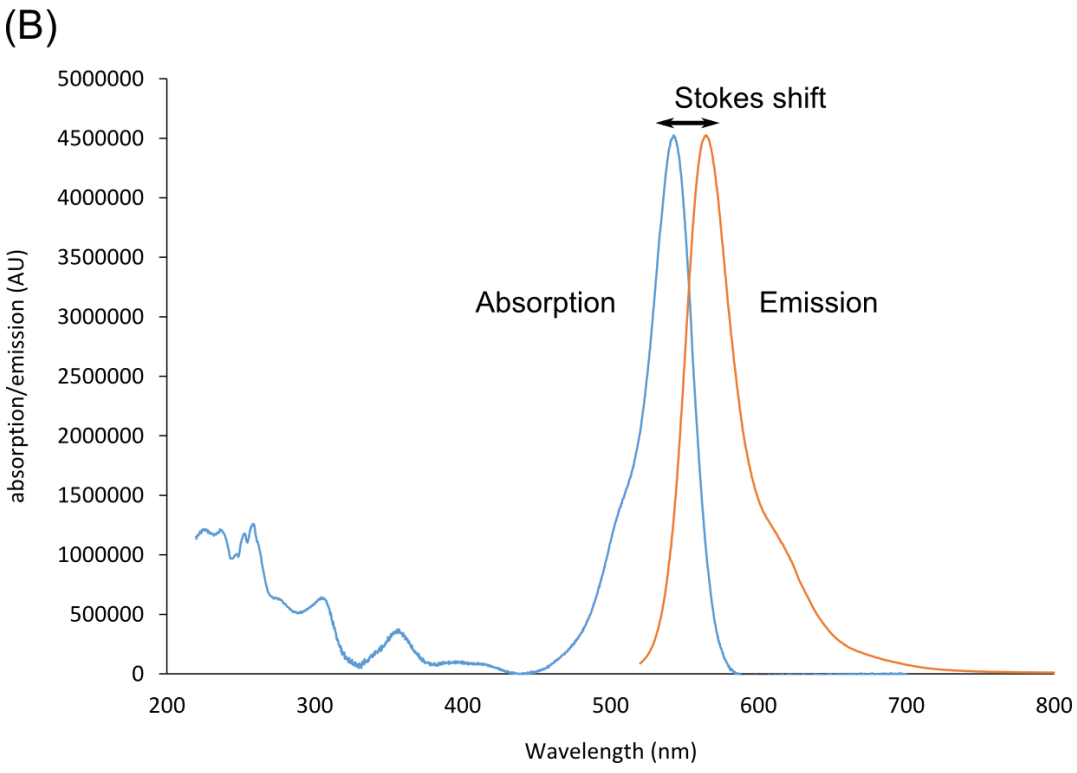
One of the big breakthroughs for instrumentation in the 20th century was the developement of fluorescence microscopy which makes use of the physical phenomenon of fluorescence to improve the contrast and therefore the resolution of microscopy. When a fluorescent molecule absorbs a photon of light electrons within the molecule are excited, they "jump" up to a higher energy state. Figure 1.1 includes a simplified Jablonski diagram illustrating this process. When the electrons relax back down to their initial energy state a photon will be emitted with a longer wavelength than that of the absorbed photon. This relaxation and fluorescence emission happens on a time scale of nanoseconds. The wavelength change between excitation and emission is known as the Stokes shift (Stokes, 1852) and occurs because vibrational relaxation (energy transfer to adjacent molecules) and internal conversion (transition between electron orbital states) lead to small losses of energy in the excited fluorophores (Lichtman and Conchello, 2005). It is also possible for intersystem crossing (where the electron spin is flipped) to result in a triplet state from which the electron cannot immediately relax, this can result in delayed fluorescence or phosphorescence on a time scale of milliseconds to hundreds of seconds. Having electrons in the triplet state can increase the reactivity of the fluorophore with the solvent, which can lead to photo-bleaching, a photo-chemical reaction that usually results in cleavage of covalent bonds within the fluorophore leading to permanent loss of fluorescence.

Figure 1.1 Excitation, fluorescence and the Stokes shift

(A) Adapted from (Johnson and Davidson, 2012). Within a molecule electrons mostly exist in a ground state (S_0), their lowest energy state. If they absorb energy, for example in the form of a photon of light, they "jump" up to a higher energy state called an excited singlet state (S_1). For the electron to relax back to the ground state the absorbed energy must be released, this can occur by a number of pathways. In fluorescent molecules the energy can be lost as a photon of light. However as energy is also lost by internal conversion the emitted fluorescence will generally be of a longer wavelength than the excitation light, this is called the Stokes shift. When an electron is in the excited state intersystem crossing (where the spin of the electron is flipped) can result in transition to a triplet state, from which it can return to the ground state via non-radiative relaxation, delayed fluorescence or phosphorescence.

(B) The absorption and emission spectra for Rhodamine B (Prahl, 2014). Absorption has been normalised such that the peak in absorption matches the peak in emission. Plotted in arbitrary units against wavelength. Rhodamine B has a maximum absorption at ~543 nm and a maximum emission at ~565 nm demonstrating the Stokes shift.


(A)
Excited Singlet states
S2
S1
S0
Internal Conversion
Vibrational Relaxation
Triplet state
T1
Intersystem Crossing
Fluorescence
10-9 - 10-7 seconds
Non-Radiative
Relaxation
Phosphorescence
10-3 - 102 seconds
Ground state


(B)
Stokes shift
Absorption
Emission
absorption/emission (AU)
Wavelength (nm)
5000000
4500000
4000000
3500000
3000000
2500000
2000000
1500000
1000000
500000
0
200
300
400
500
600
700
800

In microscopy the fact that the wavelength of fluorescence emission is shifted in relation to the wavelength of excitation means that filters can be applied which transmit only the wavelength of emission. Thus, emitted photons are separated from the excitation wavelength and the background is reduced, giving a much higher contrast. As a result of this improved contrast, use of fluorescence in biological microscopy has greatly improved the level of detail that can be studied in cells. Otto Heimstaedt and Heinrich Lehmann are credited with developing the first fluorescence microscope (1911-1913), which they used to study auto-fluorescence in bacteria, animal and plant cells (Rottenfusser *et al.*, 2016).

1.2.2.1 Labelling strategies for fluorescence microscopy

In order to study cells using fluorescence microscopy it is necessary to either find naturally occurring auto-fluorescence in the cell, as Heimstaedt and Lehmann did, or to introduce fluorescence into the cell. Examples of auto-fluorescent compounds commonly found in cells include the amino acid tryptophan (Brancaleon *et al.*, 2001), NAD(P)H (Eng *et al.*, 1989; Skala *et al.*, 2007) and compounds in the flavin family (Reinert *et al.*, 2004). However use of auto-fluorescence is limited to imaging these intrinsic fluorophores. It is now more common to use extrinsic fluorophores in order to add fluorescence to a target (which may be a tissue type, an organelle or even a specific protein). A variety of fluorescent dyes have been discovered or developed for use in biology, these can be roughly categorised as organic dyes and fluorescent proteins. Quantum dots have also been used in biological applications but this is outside of the scope of this thesis, for a review see (Jamieson *et al.*, 2007).

1.2.2.1.1 Organic fluorophores

The first organic fluorophore described was quinine sulphate, identified by Sir John Herschel, which is colourless in solution but can be excited by UV light and emits "an extremely vivid and beautiful celestial blue colour" (Herschel, 1845). Some of the older organic dyes from which many other dyes have since been derived are fluorescein, rhodamine, cyanine and boron-dipyrromethene (BODIPY). Organic dyes can have undesirable properties such as environmental sensitivity, insolubility in water, cytotoxicity and inability to permeate the membrane which can make them problematic for use in cells. However they tend to have good photo-stability and brightness in comparison to fluorescent proteins and, because they tend to be small molecules, they are less likely to cause steric hindrance to the target.

The difficulty of introducing an organic fluorophore into a cell tends to increase with increasing size. Very small, neutral charge molecules (<1 kDa) can be membrane permeable. However larger probes tend to require more invasive methods to introduce them into cells. Micro-injection uses a glass micropipette to inject a fluorescent dye directly into the cell. However this is a technically challenging process and requires that cells have a large volume and not be in suspension. Permeabilisation creates pores in the cell membrane to allow organic dye molecules to diffuse into the cell. This can be achieved by using chemicals which digest components of the cell membrane such as saponin, a plant derived glycoside (Behbehani *et al.*, 2014). It can also be achieved by osmotically shocking cells (Leidenheimer and Harris, 1991) or by electroporation which uses electrical impulses on cells in suspension to increase membrane permeability (Neumann and Rosenheck, 1972).

Besides the problems of getting organic fluorophores into cells there is also the need to target them to specific structures or cellular components of interest. A number of organic

dyes have been developed which tend to accumulate in certain organelles, which is useful for biologists. For example: DAPI (Tarnowski *et al.*, 1991) and Hoechst (Latt *et al.*, 1975) are both able to penetrate the cell membrane, bind to DNA and accumulate in the nucleus. ER-Tracker™ green and red are BOPIDY-derived fluorescent dyes that localise to the endoplasmic reticulum. However these dyes can have non-specific activity (Zünkler *et al.*, 2004) and can also have cytotoxic effects (Bielawski *et al.*, 2001).

1.2.2.1.2 Immunofluorescence

One of the major breakthroughs in biological fluorescent imaging was in the early 1940s when a scientist called Albert Coons developed a technique to label antibodies with organic fluorescent dyes and started the field of immunofluorescence (Coons *et al.*, 1941). This allowed the labelling of specific proteins with fluorescent molecules for visualisation using an antibody raised against them (Coons *et al.*, 1942). Immunofluorescence can either be direct, by labelling the specific antibody with a fluorescent dye, or indirect, by using a secondary antibody labelled with a fluorescent dye. The direct method reduces the chances of non-specific binding by using a single antibody. The indirect method has the advantage that commercially produced fluorescently labelled antibodies can be used. Multiple secondary antibodies can bind to a single primary antibody which has the advantage of amplifying the signal but makes quantification impossible. Organic fluorophores used for labelling antibodies include FITC (a fluorescein derivative), TRITC (a rhodamine derivative) and the Alexa Fluor® range. The Alexa Fluor® dyes come in a range of excitation/emission spectra derived from different starting fluorophores. They have the advantage that they have been engineered to have better stability and brightness and reduced pH sensitivity and that they are negatively charged and hydrophilic making them soluble in water. However there are significant issues with getting antibodies into living

cells due to their large size. Even if microinjected, or introduced by other means there remain issues concerning the washing out of unbound antibody and the prevention of non-specific binding. For these reasons the vast majority of immunofluorescence studies are performed in fixed permeabilised cells (Jamur and Oliver, 2010).

1.2.2.1.3 Tag labelling techniques

A more recent development in targeting inorganic fluorophores to specific proteins are tag technologies such as the HALO (Los *et al*., 2008) and SNAP (Keppler *et al*., 2002) tags. These techniques use a genetically encoded tag and a fluorescently labelled ligand that form a covalent bond when they come into contact. This allows a 1:1 labelling ratio, high specificity and, because the interaction between tag and label is covalent, excess label can be washed out without affecting the signal. However this technique is still very much dependent on the ability of the labelled ligand to penetrate the cell membrane so is subject to the same problems described above.

1.2.2.1.4 Fluorescent proteins

The discovery of jellyfish (*Aequoria victoria*) green fluorescent protein (GFP) (Shimomura *et al.,* 1962; reviewed in Shimomura, 2005) was a major step forward for biologists using fluorescence microscopy. Once the DNA sequence was known, genes encoding a chimeric fusion protein, consisting of the DNA sequence for the protein to be studied followed by a short linker sequence and then the sequence for the fluorescent protein, could be created and transformed into cells. This could be done either by inserting the labelled gene as part of a plasmid to be expressed exogenously or by replacing the wild type gene with the labelled strain. Such cells would then express the fluorescent fusion protein *in vivo*. By expressing these chimeric proteins *in vivo* the localisation of a protein within the cell could

be directly studied. Since this technique theoretically gives a precise 1:1 labelling ratio rough quantification could also be attempted by bulk studies of fluorescence intensity. Additionally, and importantly, there was no need to fix the cells to image them, opening up the option of time-lapse imaging of live cells.

Since the discovery of GFP many groups have worked on producing fluorescent proteins with different properties either through genetic modification of GFP (such as YFP, (Ormö *et al.*, 1996)) or by the isolation of novel fluorescent proteins from different species such as EosFP which was isolated from the coral *Lobophyllia hemprichii* (J. Wiedenmann *et al.*, 2004). The generation of multiple fluorophores which are excited by, and fluoresce at, distinct wavelengths of light allowed the localisation and interaction of proteins *in vivo* to be studied in unprecedented detail.

1.2.3 The limit of resolution

One of the major limitations of fluorescent microscopy rapidly became the ability to resolve distinct structures. The ability to visually separate two point-sources of light is limited by the diffractive properties of the light itself and by certain physical properties of the microscope system being used. In a widefield microscopy system this limit is described by an equation known as the Abbe limit (Abbe, 1873) (equation *1-1*).

$$\boldsymbol{r = \frac{\lambda}{2NA}} \qquad \textit{1-1}$$

Where r is the distance two point emitters need to be apart to be resolvable, λ is the wavelength of light and the numerical aperture (NA) is a physical property of the objective relating to the range of angles over which the system can collect light (Figure 1.2 A). The NA of a system can be calculated using (equation *1-2*).

$$NA = n \sin \theta \quad \textit{1-2}$$

Where n is the refractive index of the immersion medium and θ is the angle of incidence, normal to the optical axis (reviewed by Lichtman and Conchello, 2005), which is also affected by the refractive index of the immersion medium according to Snell's law (Figure 1.2 B). A higher NA objective allows light to be collected at higher angles and therefore more incident photons will reach the detector of the optical system. However, it is worth noting that a higher NA objective will also have a shallower depth of focus, something not necessarily desirable when studying a three dimensional (3D) system. The maximum NA value is limited by the technological possibilities of lens manufacture and the objectives with the highest commercially available NA have an NA of 1.4-1.49

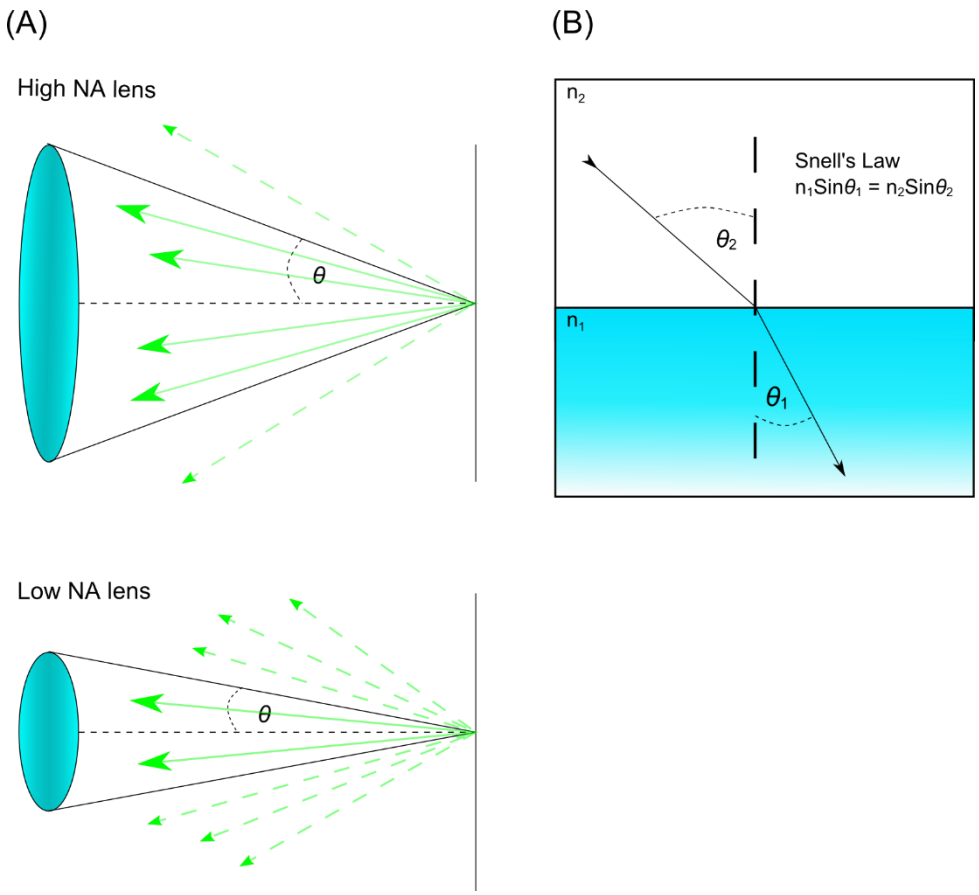


Figure 1.2 Numerical aperture and Snell's law

(A) The numerical aperture of a lens or an objective determines the angle from which it can collect light from the focal object. The higher the NA the wider an angle it can collect from and therefore the more photons it can collect per fluorophore. The solid green lines represent light that is collected and the dashed green lines represent light that is not collected.

(B) Snell's law describes how light is refracted when travelling between one medium and another. In this equation n_1 is the refractive index of medium 1 and n_2 is the refractive index of medium 2. The angle of the beam of light relative to the normal will change depending upon the angle of incidence and the change in refractive index. In a microscope system this happens when light travels from the immersion medium into the glass lens.

In a diffraction-limited system the distribution of the photons coming from a point source and forming the image is described by the point spread function (PSF) which is specific to each microscope. Within the focal plane the PSF can be approximated by a two-dimensional Gaussian function (Santos and Young, 2000) or the Airy function (Airy (Sir), 1835) (see 1.2.5.2, Figure 1.7). Outside of the focal plane the PSF has a more complex pattern.

If two point sources are closer together than the Abbe limit then their PSFs will overlap and it will be impossible to separate them visually (Figure 1.3). The practical result of this is that in conventional fluorescence microscopy the resolution limit is ~200 nm. Given that much that occurs in cells can be on the single nm - 10s of nm scale this is significantly limiting for the study of biological molecules.

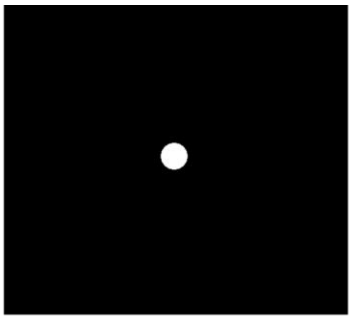
Figure 1.3 Illustrating the diffraction limit

(A) Shows a single point source of light, when this point source is visualised down the eyepiece of a microscope or on a camera it will appear as in (B). The light is diffracted giving a spread of light intensity termed the PSF where the light appears most intense in the centre of the spot and becomes more diffuse towards the edges. Any cross section through the centre of the PSF (assuming the source is in-focus) will give an approximately normal distribution of light intensity (right panel).

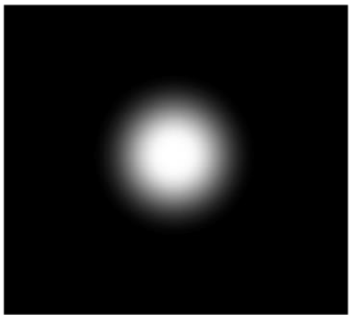
Two well-separated point sources are easily differentiated (C) but as they become closer together (D) it becomes more difficult to tell the two objects apart, until they become effectively indistinguishable (E).

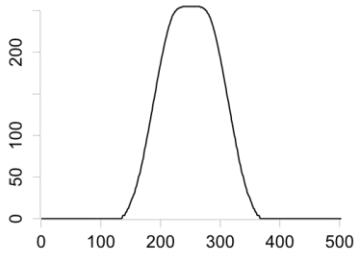
Using the Abbe limit (equation *1-1*) it is possible to calculate that the diffraction limit for a standard fluorescence microscope is around 200 nm (for example; when using 488 nm light to excite GFP with an emission peak at 509nm). Most bio-molecular processes occur at much shorter distances.

(A)

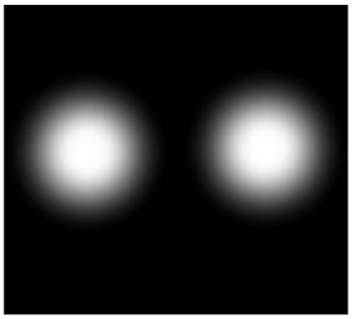


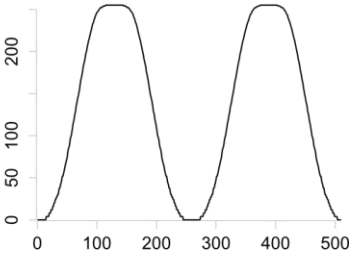
(B)



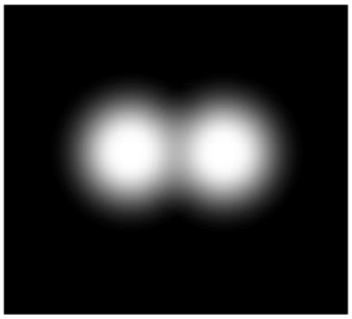

0
50
100
200
0
100
200
300
400
500

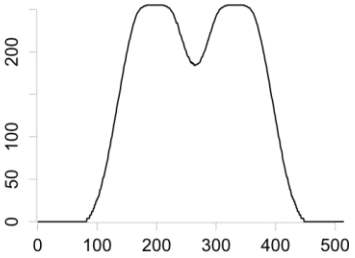
(C)



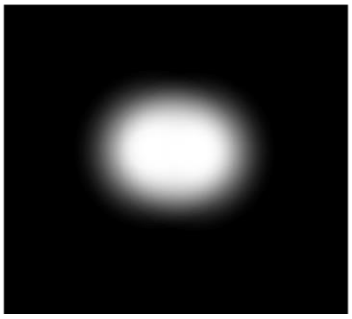

0
50
100
200
0
100
200
300
400
500

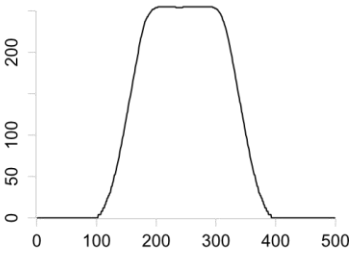
(D)




0
50
100
200
0
100
200
300
400
500

(E)




0
50
100
200
0
100
200
300
400
500

1.2.4 Illumination schemes to improve the signal and resolution in fluorescence microscopy

In fluorescence microscopy, in order to produce the best images (and thus achieve the best resolution within the diffraction limit) it is important to have a high signal-to-noise ratio (SNR) either by having a fluorophore with a very high photon yield or by imaging with a very low background, but preferably both. One of the main barriers to achieving a good SNR is the background fluorescence produced by out-of-focus fluorophores. As discussed above to achieve a good resolution it is also desirable to use a high NA objective. This limits the depth at which light is in-focus, so in a 3D cell there are likely to be many fluorophores that are outside of the focal plane.

A number of different illumination schemes exist, some of which can reduce the background created by out-of-focus fluorophores (Figure 1.4). The simplest illumination scheme is epi-fluorescence (Figure 1.4 A) in which the excitation light is aligned such that it passes straight through the objective and up through the sample. This means that the entire depth of the sample is excited to the same extent and, as such, depending on the thickness of the sample, there can be a lot of out-of-focus fluorescence.

Figure 1.4 Different illumination schemes for microscopy

Adapted from (Liu *et al.*, 2015)

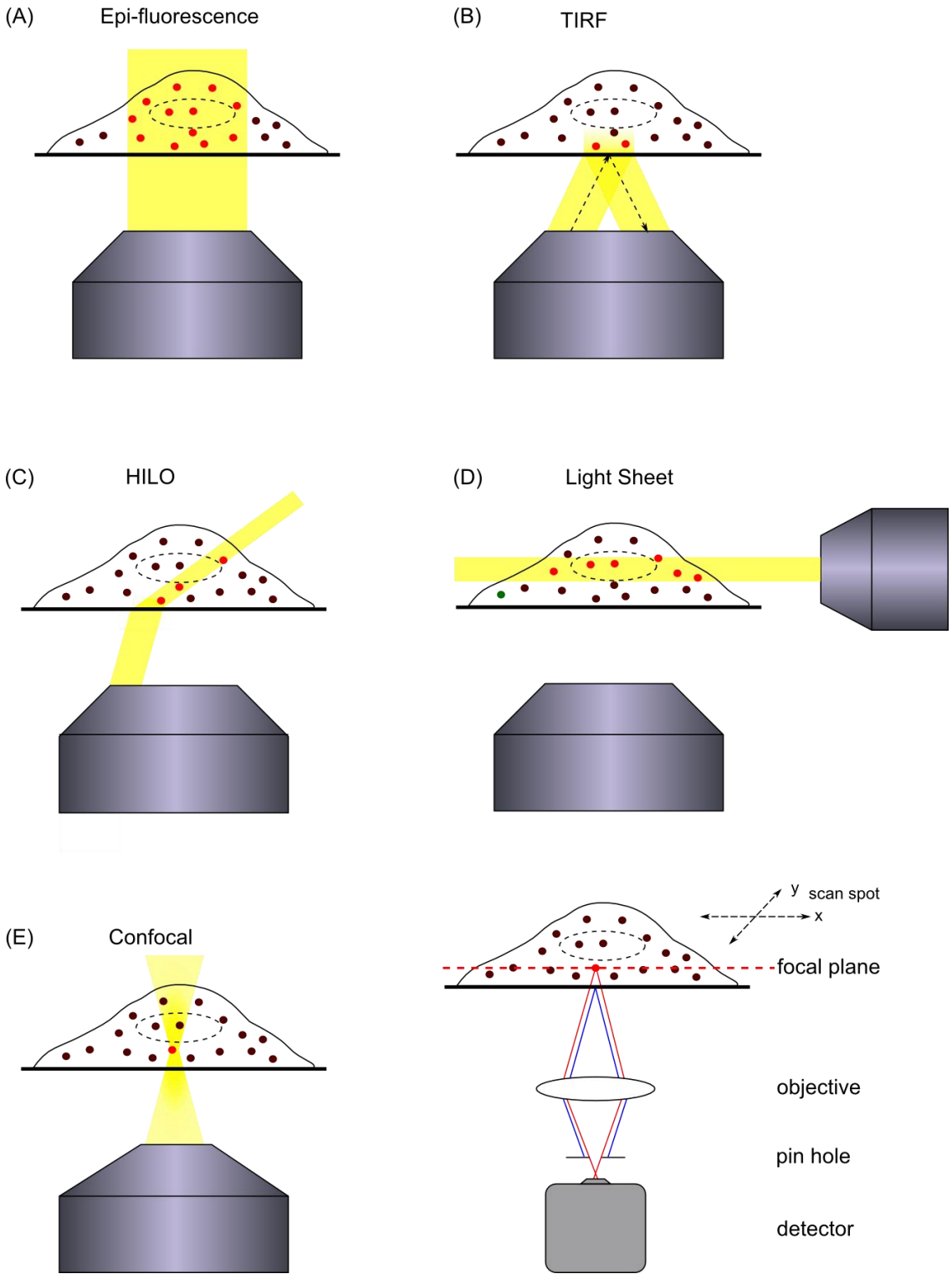
(A) Epi-fluorescence is the simplest illumination scheme. The excitation beam is aligned so it passes straight through the objective and into the sample, illuminating the full width of the sample.

(B) TIRF uses an angle of incidence of light greater than the critical angle which creates an evanescent wave of light at the interface between the glass slide and the sample. This evanescent wave only penetrates a small way into the sample, <200 nm, and only fluorophores in this volume will be excited. This effectively removes out-of-focus signal. However it can only be used to look at fluorophores on the surface of the cell.

(C) Hi-Lo illumination (also known as dirty-TIRF and oblique illumination) uses an angle not quite sufficient for TIRF but which causes the illumination profile to be narrower than that achieved by epi-fluorescence, thus reducing out-of-focus fluorescence.

(D) Light sheet illumination uses a special objective at 90 ° to the sample to create a sheet of illumination that will only excite a slice of the sample. Again this ensures that only fluorophores in the plane of focus are excited. With the right apparatus this technique can also be used for optical sectioning of a sample and creation of a 3 dimensional z-stack.

(E) Confocal microscopy focuses the excitation light to a spot in the sample and additionally uses a pinhole in front of the camera to exclude out-of-focus light (shown in blue). Only the centre of the spot has a sufficient power density to excite the fluorescent molecules, by scanning this spot across the whole sample in the x and y axes a 2 dimensional image can be drawn, by also scanning in z a stack of images can be produced which give 3D information.


(A)
Epi-fluorescence
(B)
TIRF
(C)
HILO
(D)
Light Sheet
(E)
Confocal
y
scan spot
x
focal plane
objective
pin hole
detector

One potential solution to reduce out-of-focus fluorescence is total internal reflection fluorescence (TIRF) in which the excitation beam path is aligned such that the angle of incidence is greater than the critical angle (Figure 1.4 B) (Axelrod, 1981). When this is the case at the glass-sample interface the majority of fluorescence is reflected and an evanescent wave is created which penetrates only a short distance (<200 nm) into the sample. The effect of this when imaging cells is to only excite fluorophores on, or close to, the cell surface that is in contact with the glass slide. This greatly reduces the out-of-focus fluorescence and thus increases the SNR. However this illumination scheme is only suitable for studying proteins found on, or close to, the cell membrane.

Highly inclined and laminated optical sheet (HILO) illumination (also sometimes known as dirty TIRF or oblique TIRF) uses an incident angle slightly less than that of TIRF with the effect that the beam is refracted into a highly inclined and thinned sheet passing through the sample (Figure 1.4 C). In this scenario only fluorophores in a section of the sample are excited (Tokunaga *et al.*, 2008). The benefit HILO has over TIRF is that excitation light penetrates further into the cell whilst still excluding much out-of-focus fluorescence.

A similar but more refined technique called light sheet illumination (also known as selective plane illumination) uses an objective perpendicular to the sample to produce a very thin "sheet" of light which illuminates a thin section of the sample, again reducing the interference of out-of-focus fluorescence (Huisken *et al.*, 2004) (Figure 1.4 D). Light sheet illumination has the additional bonus that, by moving the sample up and down through the light sheet, multiple sections can be imaged to build a z stack of images for 3D projection (this can be done faster than it can in confocal microscopy - see below). A disadvantage of light sheet microscopy is the physical setup for illumination: the perpendicular objective needs to be very close to the sample to maintain a thin light sheet. Additionally, on a single cell scale, light sheets created with a Gaussian beam can be too thick for proper optical

sectioning. Recent developments have partially overcome this by creating a Bessel beam and sweeping it across the sample in the *x* plane to give a scanned light sheet (Chen *et al.*, 2014; Planchon *et al.*, 2011). However this has significant costs in terms of hardware and set-up expertise.

An alternative to producing thin sections of light across a cell in order to minimise out-of-focus light is confocal microscopy (reviewed in Amos and White, 2003) (Figure 1.4 E). Confocal microscopy uses an excitation beam that is focused into the smallest possible point, meaning the power density is only sufficient to excite fluorescent molecules within the limited volume of the focus, reducing out-of-focus fluorescence. This limited excitation volume can be further enhanced by the use of two-photon excitation in which the illumination wavelength used is twice that of the optimal absorption of the fluorophore and so two photons must be absorbed simultaneously for excitation to occur (Denk *et al.*, 1990). The chance of two photons being absorbed outside of the focal volume is very low therefore the chance of fluorescence outside of the focal volume is very low. In addition placement of a pinhole in the image plane, in front of the detector, excludes out-of-focus fluorescence. The excitation volume can then be scanned across the sample in a grid pattern to build up a two dimensional image and, in addition, the sample can be scanned up or down to create optical sections at different depths, which can be used to create a 3D projection. This optical sectioning is only made possible by the exclusion of the out-of-focus light, which is achieved by minimising out-of-focus excitation due to the lower power density beyond the point of focus, combined with the use of the pinhole in the image plane. Confocal microscopy gives an improved SNR and resolution compared to widefield techniques. However the time it takes to scan the confocal point across the image means that confocal acquisitions are slow, making the technique less useful for live cell imaging. In addition, because each pixel is only illuminated for a short time, a high power density is needed in order to return sufficient photons from the fluorophore, this can cause photo-

damage to the sample, which is not conducive to live cell imaging.

Although all of these techniques help to achieve better images their resolution is still limited by the diffraction of light.

1.2.5 Breaking the diffraction limit - Super-resolution microscopy

The majority of techniques for breaking the Abbe limit in microscopy fall into two classes. The first class is the deterministic techniques, which exploit the non-linear relationship between fluorophore emission rate and illumination. These include stimulated emission depletion microscopy (STED) (Hell and Wichmann, 1994) and reversible saturable optical linear fluorescence transitions (RESOLFT) (Hofmann et al., 2005). The other class is the stochastic techniques, consisting mostly of single molecule localisation microscopy (SMLM) techniques, which stochastically separate molecules temporally and rely on image reconstruction, such as stochastic optical reconstruction microscopy (STORM) (Rust *et al.*, 2006a), photo-activated localisation microscopy (PALM) (Betzig *et al.*, 2006a) and fluorescence photo-activation localization microscopy (FPALM) (Hess *et al.*, 2006).

The SMLM techniques attempt to temporally separate molecules which are close together in space. STORM, PALM and FPALM all work on the same principle of using super-localisation of individual fluorophores. This super-localisation is achieved by actively controlling the fluorescent state of the molecule to separate individual PSFs over time. Each separate PSF can then be localised (see below) and the position of the emission centre identified with an accuracy below the Abbe limit. Fluorophores have been found whose fluorescent characteristics can be altered with a certain wavelength of light, essentially allowing them to be "switched on" so that they become excitable (this can be

reversible or irreversible). In some instances (photo-activation or photo-switching) absorption of a near-UV wavelength photon can convert a fluorophore to a second form with a specific excitation wavelength. In other instances a high-powered laser can be used to push fluorophores into a dark state. These will subsequently stochastically relax back to the ground state from which they can be excited.

In PALM and FPALM, using photo-activatable and -switchable fluorophores, if a sufficiently low power density is used for the activating light then only a small subset of the fluorophores will be switched. If the subset is sufficiently small they should be spatially separated. These activated fluorophores can then be excited and imaged as in conventional fluorescence microscopy. Eventually, continuous excitation leads to photo-bleaching, whereby the fluorophore undergoes a photo-chemical reaction, usually cleavage of covalent bonds, which leads to permanent loss of fluorescence. By exciting the activated subset for a suitable period all the active fluorophores should be photo-bleached, a process that is irreversible, and therefore the fluorophore will not be imaged again. Multiple rounds of activation and image capture can be performed iteratively, until all the fluorophores within the field of view have been activated, visualised and photo-bleached. Software (such as downloadable packages like QuickPALM or RapidSTORM, or custom written software) is then used to calculate the position of the molecule based on a model of the PSF. The locations is predicted by matching the model of the PSF to the image (see 1.2.5.2). This enables reconstruction of a high resolution image from the data (Figure 1.5).

STORM follows the same principle of temporal separation of PSFs but differs a little from PALM and FPALM. In STORM reversibly photo-switchable proteins are used which can be pushed into a dark state. In the original STORM paper a Cy3-Cy5 dye pair was used (Rust *et al.*, 2006a). The Cy5 would fluoresce under low powers of a 633 nm laser and under higher powers would be pushed into a dark state. In combination with the Cy3 a 532

nm laser could be used to return (or switch) the Cy5 to the fluorescent state. By using a low power of the green laser to stochastically return only a subset to the fluorescent state the fluorophores should be well-separated and their PSFs distinguishable. The imaged fluorophores can then be pushed back into the dark state and a new subset switched. This cycle can be repeated many times before photo-bleaching occurs. In addition Heilemann *et al.*, demonstrated that what they called direct STORM (dSTORM) could be performed using single conventional fluorescence microscopy probes, rather than probe pairs, by using a higher power of green (514 nm) laser light combined with a buffer containing oxygen scavengers (Heilemann *et al.*, 2008). Since each fluorophore can be image multiple times quantitative information is lost by the repeat localisation of the same fluorophore.

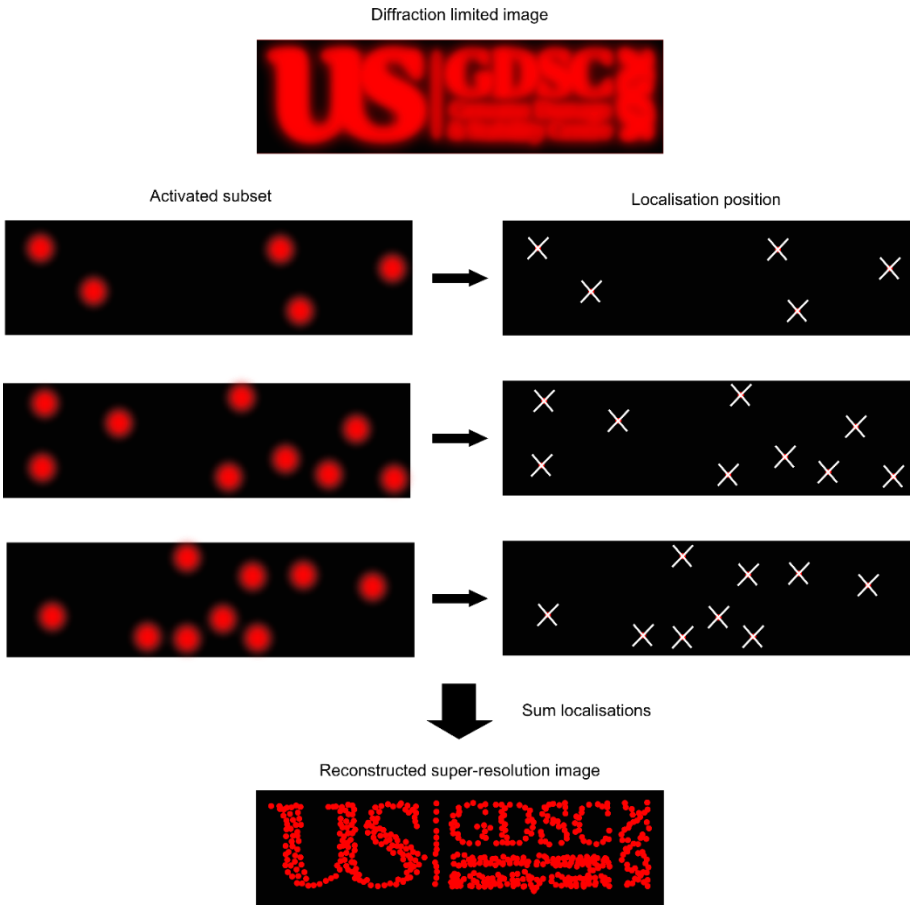


Figure 1.5 Sequential illustration of SMLM techniques

In a diffraction limited system the resolution is limited to 200-300 nm meaning that objects close together cannot be differentiated and images lack definition. SMLM techniques utilise fluorophores with controllable excitation states. These fluorophores can be activated or switched into an excitable state, usually by exposure to near UV light. Once in this excitable state they can be excited by a specific wavelength of light and will emit fluorescence which can be detected on a camera. Fluorophores excited for a long time or with a high power of laser will eventually become photo-bleached, meaning they will irreversibly lose the ability to be excited and will not be visualised again. A subset of the fluorophores present is switched/activated and if this subset is sufficiently small the chances of overlap are reduced to practically nil. An image of this subset can be saved and used to estimate the position of the fluorophores in a process called localisation. The localisation position is recorded and a different subset activated and localised. This can be repeated until all the fluorophores have been photo-bleached and no more activation is seen. Then the localisations can be summed together and re-plotted to give a higher resolution image.

1.2.5.1 Photo-switchable and -activatable fluorophores

SMLM techniques depend on the controllable fluorescent states of photo-switchable and -activatable fluorophores. Conventional fluorophores can be divided into organic dyes and fluorescent proteins and, likewise, photo-switchable (PS)/photo-activatable (PA) fluorophores can be found in both groups. Organic PA/PS dyes have the advantage of tending to be brighter, with a higher photon budget, which means they can be localised with greater precision. The disadvantages of organic dyes are that they can be harder to introduce into cells, usually they need to be coupled to antibodies, and it can be difficult to ensure a 1:1 fluorophore to target ratio (see 1.2.2.1.1). In contrast photo-activatable fluorescent proteins (PA-FPs) can be genetically encoded into cells as a chimeric fusion with the target protein. Thus, theoretically, giving a precise 1:1 labelling ratio, without non-specific binding, and circumventing the problem of introducing the fluorophore into the cell. However PA-FPs tend to be dimmer than organic dyes and can have undesirable oligomerisation properties.

Photo-activatable green fluorescent protein (PA-GFP) was the first PA-FP described, having been engineered from wild type GFP to enhance the feature of going into a dim or dark state on irradiation with 490 nm light, from which it could be returned by irradiation with 400 nm light (Patterson, 2002). Several novel PA-FPs have been isolated from sea creatures (as GFP originally was) including EosFP (Wiedenmann *et al.*, 2004) and Kaede which was isolated from the stony coral *Trachyphyllia geoffroyi* (Ando *et al.*, 2002). More recently there have been a number of attempts to engineer photo-activatable or -switchable proteins based on the structure and sequences of known proteins in an attempt to introduce more ideal qualities. For example multiple variants of EosFP have been developed which change its photo-physical (Adam *et al.*, 2008; Fuchs *et al.*, 2010) or oligomerisation properties (McKinney *et al.*, 2009; M. Zhang *et al.*, 2012). For a table of

properties of PA-FPs see Appendix, for a review on PA-FPs see Adam, (2014).

One of the most currently popular PA-FP families is derived from Eos, a protein isolated from *Lobophyllia hemprichii* (J. Wiedenmann *et al.*, 2004). The isolated protein was found to absorb at a wavelength of ~503 nm and fluoresce with an emission wavelength of ~516 nm. However, after it had been exposed to near UV light (~400 nm) it would shift to a red state which would absorb at ~571 nm and emit at ~581 nm (Figure 1.6). The green form of Eos has a cis-coplanar two ring chromophore. The red form is generated by cleavage of the peptide backbone at the His^{62} α carbon (Cα) which leads to the formation of a double bond between the His^{62} Cα and Cβ, this extends the π conjugation to the His^{62} imidazole ring (Mizuno *et al.*, 2003). Computer simulations have been used to suggest that photo-conversion occurs when the green form is excited to the S_1 singlet state and undergoes intersystem crossing (ISC) to the triplet state (see Jablonski diagram in Figure 1.1) from which peptide bond cleavage proceeds (Lelimousin *et al.*, 2009).

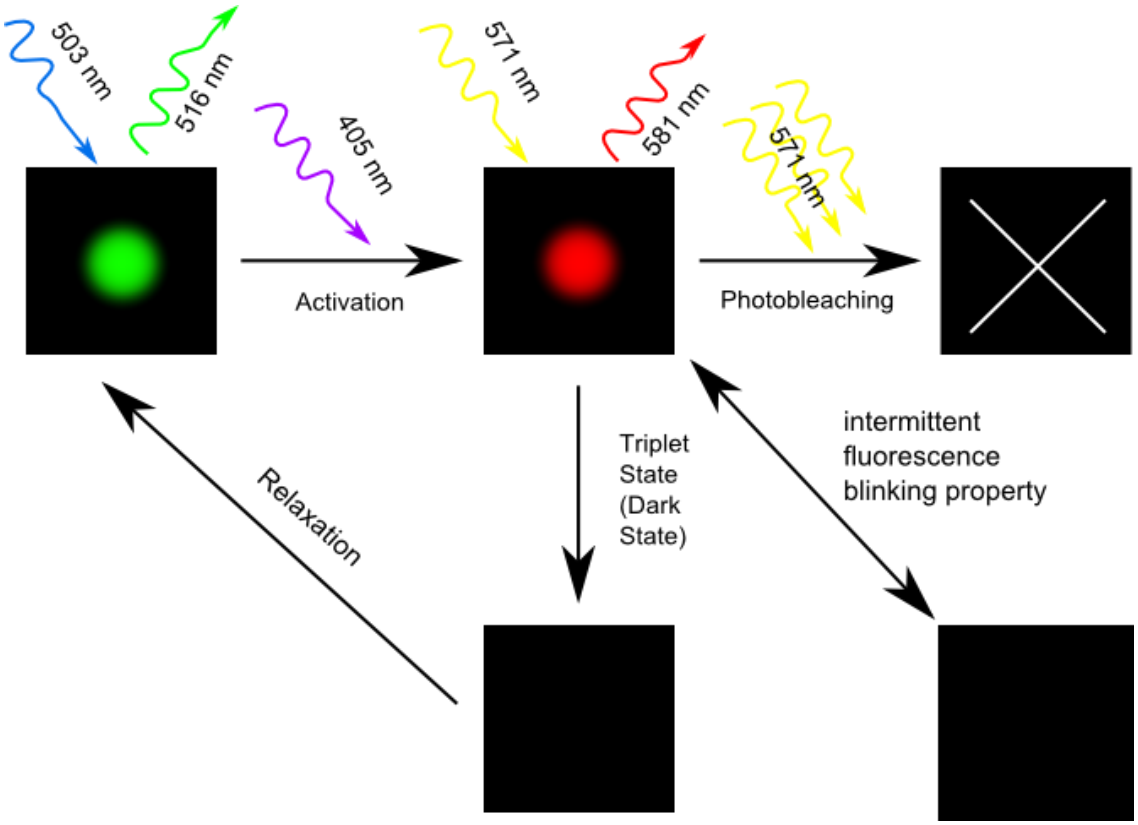


Figure 1.6 The Fluorescence States of the Eos photo-activatable fluorescent protein

In its base state Eos and its derivatives exist in a green form. This form is excited optimally by light around ~503 nm and emits at ~516 nm. On irradiation with near UV light (405 nm) the fluorophore undergoes a change and starts to absorb optimally at ~571 nm and emit at ~581 nm, this is the red form. In addition, when the red form is excited it does not emit constantly but rather shifts between the excited state and a short lived dark state leading to a fluorescence intermittency property that is known as blinking. This is a common property of photo-switchable dyes and can often be induced in other fluorophores. Further the Eos proteins appear to have a second, long-lived, dark state, probably a triplet energy state. When excitation with the 571 nm light is stopped the fluorophores appear to be able to relax out of this dark state and become re-activatable by 405 nm light.

The wild type EosFP protein exists as a tetramer in solution which makes it problematic as a protein tag, as tetramerisation may result in destabilisation of the protein of interest or mis-localisation. Several mutants have been produced attempting to create a monomeric form whilst maintaining the spectral properties of the original. One of the more apparently successful was mEos2 (McKinney *et al.*, 2009) which contained mutations reported to make it monomeric, whilst remaining very bright and with a good photon yield. Later studies, including purifications performed in the Carr laboratory of the Genome Damage and Stability Centre (GDSC) at the University of Sussex, have demonstrated that mEos2 can form a multimer in solution, which could interfere with the functionality of tagged proteins or their ability to oligomerise normally (Dr Adam Watson, unpublished). Other problems with mEos2 include a short-lived reversible fluorescence intermittency, termed 'blinking', and a long-lived dark state (Figure 1.6). The 'blinking' means that once mEos2 is photo-converted to its red form and excited it is not continuously emitting until it is photo-bleached, but rather appears to switch between a fluorescent 'on' and 'off' state (Annibale *et al.*, 2010). In quantitative studies this blinking property can lead to over counting as it is difficult to ascertain the difference between a newly activated fluorophore and one that has temporarily undergone a blinking event. This can be dealt with to some extent by using "pulsed" activation, having a long period between activation pulses and assuming that any fluorescent spots localised to the same position within the pulse window come from the same molecule. The dark state is more challenging because it appears that in some instances, instead of being photo-bleached, mEos2 can enter a dark state from which it can be re-activated, preventing it from being photo-bleached and thus allowing the same molecule to be counted multiple times over a long time period.

In 2012 Zhang *et al.*, demonstrated that mEos2 also formed oligomers in solution and they came up with two alternatives that they believed to retain the brightness of mEos2 whilst being truly monomeric. These were named mEos3.1 and mEos3.2. Zhang *et al.* looked into

potential sites for oligomeric binding and used site directed mutagenesis to attempt to disrupt this. The mutations made to mEos2 to produce mEos3.1 are I102N, I157V, H158E and Y189A (M. Zhang *et al.*, 2012) (see Appendix for DNA and protein sequences of mEos2, 3.1 and 3.2). The new fluorophores mEos3.1 and mEos3.2 appeared to be as bright as mEos2 and in the Carr lab they were also observed to be genuinely monomeric, but they also appear to have a lower photo-activation efficiency than mEos2. A recent study gave mEos2 a photo-activation efficiency of 0.6 and mEos3.2 a photo-activation efficiency of 0.4 (Durisic *et al.*, 2014). For the majority of work for this thesis either mEos3.1 or mEos3.2 have been used as the PA-FP tag.

1.2.5.2 Localising fluorophores from their point spread functions

The aim of localisation is to predict the position of the fluorophore based on the image of the PSF as accurately as possible. In addition it can be useful to predict other properties of the PSF such as width or number of photons. While there are a number of methods that predict the position but do not give any information about any of these other properties, for example centre-of-mass localisation, the principle methods for localisation work by attempting to match the data to a model of the PSF. There are two well described models for the PSF in the focal plane; the Airy pattern and the Gaussian distribution. Outside of the focal plane the PSF differs from a spot due to interference patterns. Moving away from the focal plane the PSF becomes increasingly broad and dim meaning that either the spot will not be localised or it will be localised with poor precision and can be discarded on the basis of width. The Airy pattern, which takes into account the fact that the waves of focused light create an interference pattern of concentric rings around a central point (Airy (Sir), 1835) is the more complicated model. However in a real optical system it is quite likely that the PSF will differ from the Airy pattern as a result of optical aberrations and

other properties of the system. When the Gaussian distribution and Airy patterns are compared (Figure 1.7) the major differences between them occur at regions of very low signal and so are likely to be swamped by noise in real systems. As a result and because the Gaussian distribution is easier to calculate and has analytical gradient functions it is often preferred over more complex models of the PSF such as the Airy pattern.

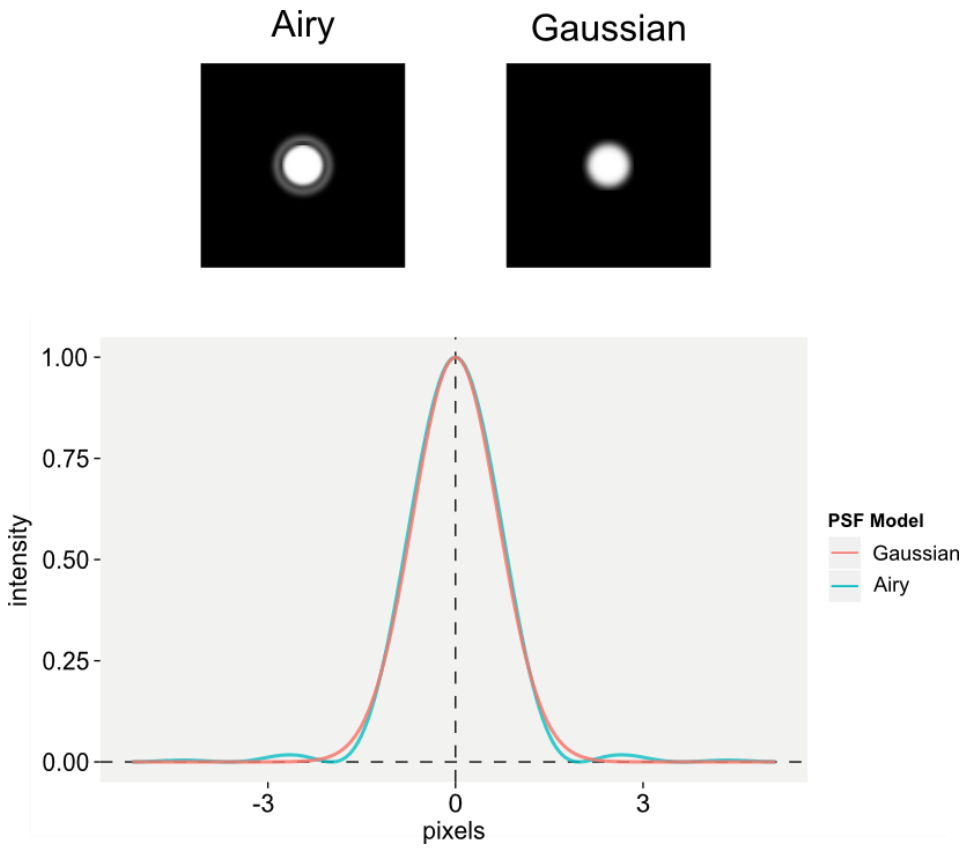


Figure 1.7 Models of the PSF for the microscope set up used

For the pixels pitch, magnification, wavelength and NA used in the microscope set up this is a model of the PSF based on an Airy function (blue) and a Gaussian distribution (red). The equation used to produce the Airy diffraction pattern is:

$$I(\theta) = I_0\left(\frac{2J_1(ka\sin\theta)}{ka\sin\theta}\right)^2 = I_0\left(\frac{2J_1(x)}{x}\right)^2$$

Where I_0 is the maximum intensity of the pattern at the centre of the disc, J_1 is the Bessel function, k is the wave number ($2\pi/\lambda$), a is the radius of the aperture, θ is the angle of observation and therefore x is the distance from the centre. The Gaussian pattern is given by:

$$I(x) = I_0\left(\exp(-0.5\left(\frac{x}{f}\right)^2\right)$$

Where f is a scaling factor set at 1.323 (Abraham *et al.*, 2009)

Having established a model there are also two main ways of fitting that model to the raw data:

The first is least squares estimation (LSE). LSE methods use a model to predict an image that is compared to the actual data; the model is then updated until no more improvements can be made. LSE methods do not require anything other than a function to produce a model PSF. The name refers to the fact that these methods attempt to minimise the sum of squares between the actual data and the predicted data. The problem with using squares is that single pixels that are far from the model can become heavily weighted, whereas ideally all pixels should be equally weighted. It is possible to have a weighted LSE where a weight is assigned to each pixel using a weighting function, for example a model of the noise at each pixel, so that noisier pixels are weighted less.

The second type of method is called maximum likelihood estimation (MLE) and calculates the probability that the actual data could be generated with the PSF model; the model is then updated to find the model with the best probability. These methods do not suffer from bias to outlier pixels but they do require a probability model for the data. This requires modelling the probability of light (generally modelled as a Poisson distribution) and of the random noise introduced by the camera (for which the EM-gain is generally modelled as a gamma distribution and the read-noise as a Gaussian).

Both these methods have been benchmarked by Abraham *et al.*, (2009) who found that MLE was more accurate in estimating data simulated and fitted using both an Airy pattern and a Gaussian distribution. Under ideal conditions MLE approached the limit of accuracy as described by the Fisher-information matrix (equation *1-4*) and the Cramer-Rao lower bound (equation *1-5*), which is the theoretical limit of precision given the noise in the data. This accuracy advantage was lost at high noise levels. However, MLE was also less sensitive to the relative location of the point source within the pixel (whether it was located centrally

or towards one edge) and to misspecification of the width parameter by the model (i.e. when using a model PSF that had an incorrect width the location was estimated with better precision using MLE than LSE). Overall they determined MLE to be the more robust fitting method.

In 2002 Thompson *et al.*, were the first to derive an equation to describe the precision of localisation fitting over a range of conditions (Thompson *et al.*, 2002). For an error in localisation (Δx) the parameters are σ, the standard deviation of the PSF, a, the size of pixels, N, the number of photons collected and b^2, the background noise.

$$\langle(\Delta x)^2\rangle = \frac{\sigma^2 + a^2/12}{N} + \frac{8\pi s^4 b^2}{a^2 N^2} \qquad 1\text{-}3$$

The theoretical limit of localisation accuracy can be described by the Fisher-information matrix and the Cramer-Rao lower bound. The Fisher-information matrix gives the amount of information that an observable variable X gives about an unknown parameter θ that the probability of X is dependent on. The function $f(X; \theta)$ is the probability function for X and also the likelihood function of θ. Where θ is some parameter of the PSF, for example a position coordinate, and E is the expected value.

$$I(\theta) = E\left[\frac{\delta}{\delta\theta}(\ln f\,(X;\theta))^2|\theta\right] \qquad 1\text{-}4$$

The Cramer-Rao lower bound states that the inverse of the Fisher-information ($I(\theta)$) is a lower bound on the variance of any unbiased estimator of θ:

$$Var(\theta) \geq \frac{1}{I(\theta)} \qquad 1\text{-}5$$

In 2010 Mortensen *et al.*, suggested three alternative formulas for localisation error for both LSE and MLE fitting with a 2D Gaussian and a poissonian noise distribution. These were: an error formula for the Gaussian mask estimator (GME) (Thompson *et al.*, 2002) (*1-6*), an error formula for the weighted least-squares Gaussian fit (WLS) (Bobroff, 1986)

(*1-7*) and an error formula for their own MLE with the theoretical PSF (MLEwT) (*1-8*).

$$Var(\hat{\mu}_x) = F\frac{\sigma_a^2}{N}\left(\frac{16}{9} + \frac{8\pi\sigma_a^2 b^2}{N^* a^2}\right) \quad 1\text{-}6$$

$$Var(\hat{\mu}_x) = Var(\hat{\mu}_y) = F\frac{\sigma_a^2}{N}\left(\frac{1}{I_1(\tau)} - \frac{2\pi\sigma_a^2 b_{offset}^2}{Na^2}\frac{I_2(\tau)}{I_1(\tau)^2}\right) \quad 1\text{-}7$$

$$Var(\hat{\mu}_x) = Var(\hat{\mu}_y) = F\frac{\sigma_a^2}{N}\frac{1}{I_1(\tau)} = \frac{\sigma_a^2}{N}\left(1 + \int_0^1 \frac{\ln t}{1 + t/\tau}dt\right)^{-1} \quad 1\text{-}8$$

Where μ_x and μ_y are the mean positions in the x and y directions, F is a factor representing the EM-CCD noise (usually 2), a is the pixel size, N is the number of photons collected, b^2_{offset} is the pixel output not due to photons recorded, $\tau = 2\pi\sigma_a^2(b^2+b^2_{offset})/(Na^2)$, σ^2 is the fitted variance of the PSF, σ_a^2 is the adjusted standard deviation of the Gaussian, b^2 is the expected number of background photons per pixel, $I_1(\tau) = -\int_0^1 \frac{t\ln t}{t+\tau}dt$ and

$I_2(\tau) = -\int_0^1 \frac{t\ln t}{(t+\tau)^2}dt$.

The model used in PeakFit (the localisation software developed by Dr Alex Herbert as part of the larger SMLM project) uses MLE fitting with a Poisson distribution for the probability of light, a Gamma distribution for the EM-gain and a Gaussian distribution for the read noise of the camera. As such it is referred to as a Poisson-Gamma-Gaussian model. It is also possible to use an LSE fit which is faster and has no requirement to know the bias, gain and read noise for the camera. It is then possible to use the Mortensen formulas for both LSE and MLE fitting as precision estimates that are quite good, but it should be noted that they assume a Gaussian PSF model is fit to a Gaussian PSF rather than a true microscope PSF.

1.2.5.3 Applications of SMLM in biological systems

Since the SMLM techniques were first described, in 2006 (Betzig *et al.*, 2006a; Hess *et al.*,

2006; Rust *et al.*, 2006a), they have been adopted by a number of research groups for use in studying biological systems. In particular, the availability of pre-fabricated microscope systems has increased the number of people trying to apply these techniques to their research. Hess *et al.*, did not use biological examples to initially illustrate their description of FPALM (Hess *et al.*, 2006) however they later used it to achieve 40 nm resolution imaging of hemagglutinin in both live and fixed cells (Hess *et al.*, 2007). Rust *et al.*, used STORM to image RecA coated circular plasmid DNA *in vitro* (Rust *et al.*, 2006a). Betzig *et al.*, were able to use PALM to image the protein vinculin at cell focal adhesions, actin within a lamellipodium and the distribution of the retroviral protein Gag at the plasma membrane *in vivo* in fixed cells (Betzig *et al.*, 2006a).

The SMLM techniques are most easily applied to fixed cells because the timescale of SMLM experiments is relatively long and, in unfixed cells, the positions of structures being studied would be likely to shift over the course of the experiment. PALM has been used in fixed *E. coli* cells to investigate the distribution of proteins needed for chemotaxis (Greenfield *et al.*, 2009). STORM has been used in Rat hippocampal neuron cells to show that actin forms rings along the axonal shafts but also longitudinal filaments deeper into the shaft (Xu *et al.*, 2013; Zhong *et al.*, 2014). It has also been used to study how the actin skeleton changes in HeLa cells infected with *Salmonella typhimurium* (Han *et al.*, 2014). Looking in mouse tumour cells Erami *at al.*, used STORM to demonstrate that E-cadherin form sub-diffraction sized clusters and was not evenly distributed through the cell (Erami *et al.*, 2015; Wu *et al.*, 2015).

Despite the time limitation many live-cell SMLM experiments have been performed: Shroff *et al.*, (2008) used PALM to observe the dynamics of proteins in adhesion complexes, transmembrane cytoskeleton-substrate attachment points needed for cell migration, by using extremely short exposure times. Manley *et al.*, (2008) also studied the dynamics of

proteins in live cells by combining PALM with single-particle tracking to study the distribution and mobility of two virus membrane proteins in the plasma membrane. Li *et al.*, have used single-particle tracking SMLM in live cells to study the dynamics of Huntingtin protein fragments, identifying three different dynamic states (Li *et al.*, 2016). Biteen *et al.*, used live cell SMLM to study the diffusion kinetics of a histidine protein kinase in *Caulobacter crescentus* cells (Biteen *et al.*, 2008; Biteen and Moerner, 2010). Recently Etheridge *et al.*, (2014) described a new technique to utilise the blurring of mobile proteins imaged at long exposure times in live cells to exclude them from imaging and thus gain information on the proportion of the protein that is static (implying association or binding to a relatively immobile structure such as chromatin) under different conditions.

There is also much interest in the quantitative potential for PALM techniques. In theory each fluorescently labelled molecule is activated and imaged only once, meaning that, assuming a 1:1 labelling ratio, it should be possible to count every fluorescently labelled protein. In 2011 a technique was developed for extracting information on the clustering of proteins in the plasma membrane, including cluster size, density and their abundance in the plasma membrane using pair-wise correlation with PALM (PC-PALM) (Sengupta *et al.*, 2011). This technique was applied by Puchner *et al.*, (2013) to study proteins in vesicle membranes in order to track their maturation. They also used it to look at artificial oligomers of mEos2 in membranes. Other uses of quantitative PALM to study membrane proteins include a study on the oligomerisation of the asialoglycoprotein in rat cells (Renz *et al.*, 2012) and a study on the formation of Raf (a kinase involved in cell growth regulation) multimers (Nan *et al.*, 2013). Quantitative studies in the membrane benefit from requiring only a narrow focal depth and as such can be used with the TIRF illumination scheme (see 1.2.4), which eliminates out-of-focus fluorescence and thus improves SNR. Quantitative studies have also been performed in the cytoplasm, looking at proteins that form a single focus in the cell (Lando *et al.*, 2012), in which case out-of-focus light and

noise can be excluded based on spatial position, and in the nucleus to give a comparative quantification of nucleosome clustering (Ricci *et al.*, 2015). The potential of quantitative PALM is in part limited by the precision with which fluorophores can be localised. The brighter the fluorophore used the better the precision. One recent study was able to quantify the number of several proteins in the nuclear pore complex of mammalian cells by use of STORM imaging and immunostaining with a very bright organic dye giving a resolution in the single nm range (Szymborska *et al.*, 2013).

1.3 Background to the biological questions

The Carr research group works primarily on investigating the mechanisms of DNA replication and the sensing and repair of DNA damage using the model organism *Schizosaccharomyces pombe*. The aim of this thesis is to understand the limitations of PALM in reference to biological problems relevant to the research currently being undertaken within the group. To this end, all biological work for this thesis has been carried out using *S. pombe* and, therefore, this introduction uses the *S. pombe* nomenclature. Within this thesis PALM has been applied to two different proteins. The first is the large subunit of the ribonucleotide reductase (RNR) complex, which is highly abundant in the cytoplasm of the *S. pombe* cell and is thought to exist in more than one stoichiometry (Nestoras *et al.*, 2010). The second is a recently described protein called Rrp2, which is thought to function in the homologous recombination (HR) DNA damage repair pathway and is expressed at such low levels it is difficult to visualise with conventional fluorescence microscopy (Dziadkowiec et al., 2009, 2013). Thus, here is presented a brief description of *S. pombe* as a model organism, a description of RNR complexes and what is known about their stoichiometry and an introduction to the *S. pombe* DNA damage response.

1.3.1 The yeast *Schizosaccharomyces pombe* as a model organism

S. pombe, also known as fission yeast, is a single celled eukaryote first isolated by Paul Linder from beer brewed in East Africa. It has a rod shape 3-4 μm in width and 8-15 μm in length. *S. pombe* only grows at its tips, and does not change diameter during growth meaning that its length increases as it progresses in the cell cycle. When it divides, it splits down the middle to give two equal sized daughter cells (Figure 1.8) making it a good model organism for studying the cell cycle. The mitotic (or vegetative) cell cycle consists of four phases; G1, which is very short in *S. pombe*, S, which lasts approximately 20 minutes and during which replication occurs, G2, which lasts several hours, and then M, during which the chromosomes segregate and the nucleus divides. Cytokinesis and septation are co-incident in the S phase. It takes *S. pombe* between 2 and 4 hours to complete a single cell cycle, dependent on growth conditions, with the majority of that time spent in the G2 phase. *S. pombe* cells have a simple genome consisting of ~13 Mb across three chromosomes (Kohli *et al.*, 1977), which has been fully sequenced (Wood *et al.*, 2002), making them an excellent tool for genetic experiments. *S. pombe* cells are amenable to genetic manipulation in that they are able to maintain autonomous plasmids and will recombine homologous sequences of linear DNA into the genome. *S. pombe* is a haploid type cell meaning gene function can be inactivated by a single mutation. Whilst *S. pombe* cells are typically haploid, in the lab diploid cell populations can be maintained if needed. In addition, haploid *S. pombe* cells can be induced to conjugate and undergo meiotic division. This will produce four daughter cells with a random assortment of the parental genes, a feature that can be used to combine mutations into a single strain. Many of the genes that encode basic cellular processes in eukaryotes are highly conserved between *S. pombe* and other organisms, and thus understanding gained from experiment with *S. pombe* cells is relevant to the

understanding of similar process in other cell types, including human. As an example of this, in 2001 the Nobel Prize in physiology or medicine was won by Paul Nurse (shared with Tim Hunt and Lee Hartwell) for his work on elucidating the cell cycle. He identified the key regulators of cell cycle control using *S. pombe* and showed that this was conserved from yeast to man by identifying human CDK1 using functional complementation of the yeast mutant

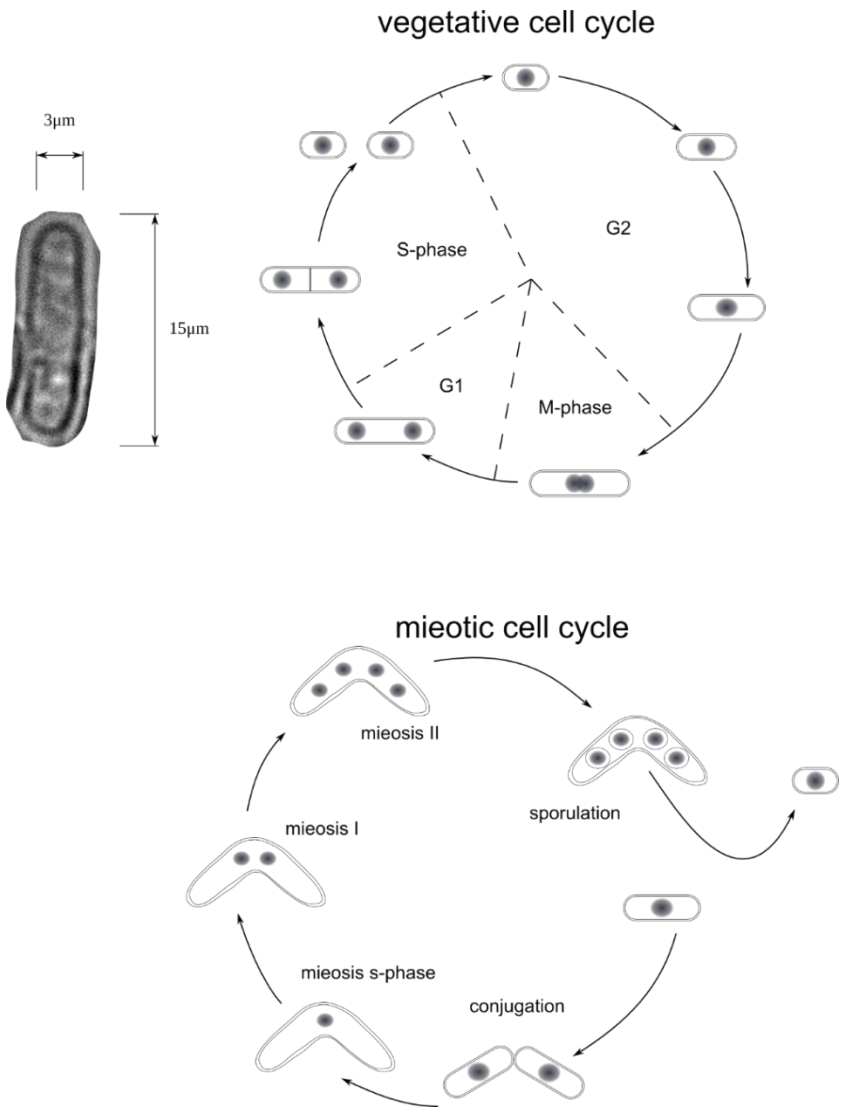


Figure 1.8 The life cycle of *Schizosaccharomyces pombe*

Adapted from (Forsburg, 2015)

S. pombe is a rod shaped cell 3-4 µm in width and 8-15 µm in length. As the cell progresses through the cell cycle it elongates without changing width, making it simple to tell the position of the cell in the cell cycle of cells. Under normal circumstances *S. pombe* cells undergo a vegetative cell cycle in which the genome is duplicated and the cell divides exactly in half down the centre to give two identical daughter cells. However *S. pombe* strains in nutrient deficient circumstances, particularly on low nitrogen media, will conjugate in order to undergo meiosis and recombine genetic material.

1.3.2 The class Ia ribonucleotide reductase complex

The RNR complex catalyses the reaction converting ribonucleotide diphosphates (NDPs) to deoxy-ribonucleotide diphosphates (dNDPs) which are the precursors to deoxy-ribonucleotide triphosphates (dNTPs). This reaction is central to a cell's ability to replicate DNA and to repair DNA damage. In eukaryotes four different deoxy-ribonucleotide monophosphates (dNMPs) are present in DNA, deoxy-adenosine-, deoxy-guanosine-, deoxy-thymidine- and deoxy-cytidine- monophosphate (dAMP, dGMP, dTMP and dCMP). The four equivalent triphosphates (dATP, dGTP, dTTP and dCTP) are therefore the necessary building blocks for DNA replication. Imbalances in the available pools of these dNTPs can lead to mutations and other genotoxic events (reviewed in Mathews, 2006). Hence, dNTP levels are tightly regulated. RNR complexes can be separated into classes I to III depending on the free radical and metal co-factor they use for reduction (reviewed in Reichard, 1993).

Eukaryotes, and some eubacteria and viruses, use class Ia RNR complexes which are typically made up of two subunits designated R1 (the large catalytic subunit, sometimes knowns as α) and R2 (the small regulatory subunit, sometimes known as β) (Nordlund and Reichard, 2006). The R1 subunit contains the active site and two allosteric binding sites termed the activity and specificity sites (Eriksson *et al.*, 1979). The activity site only binds ATP and dATP; dATP acts as an inhibitor of the RNR complex, feeding back that there is an abundance of dNTPS, conversely ATP acts as an activator. The specificity site can bind any of ATP, dATP, dGTP and TTP causing a complex pattern of promotion and inhibition of substrate selection by conformational change (

Figure 1.9) (reviewed by Nordlund and Reichard, 2006). The R2 subunit contains a tyrosyl free-radical which is used to catalyse the reduction reaction (reviewed in Gräslund *et al.*, 1985). In *S. pombe* the R1 subunit is called $Cdc22^{R1}$ and the R2 subunit is called $Suc22^{R2}$

(Sarabia *et al*., 1993).

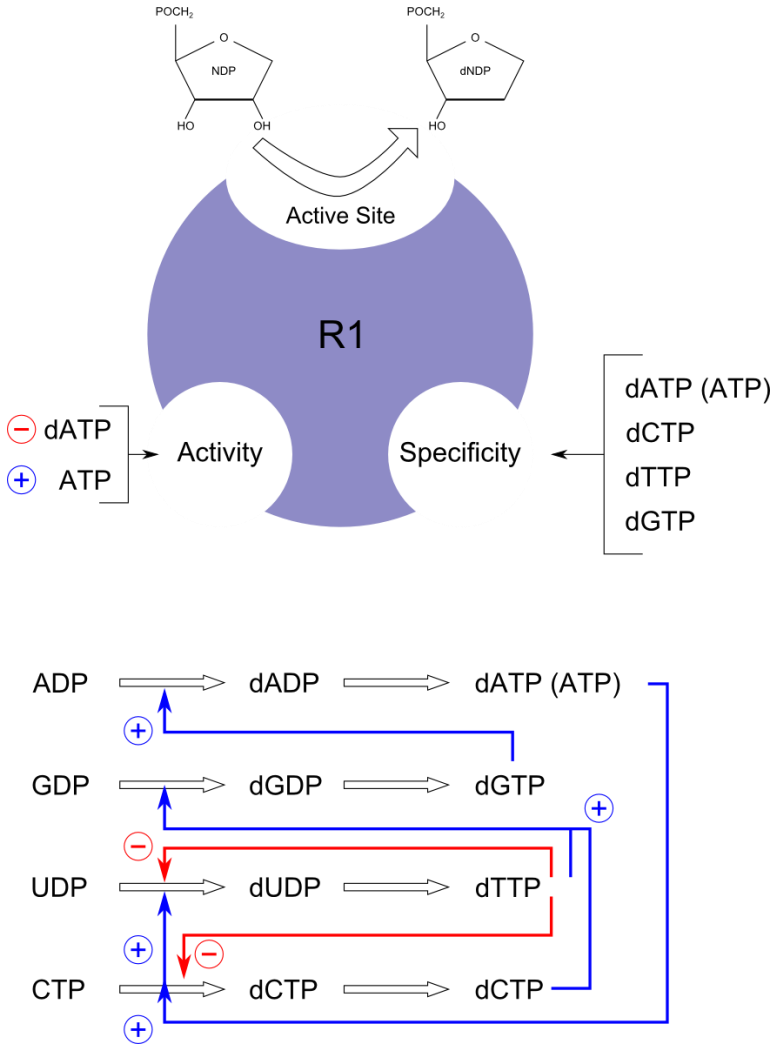


Figure 1.9 Allosteric regulation of the RNR complex

Adapted from (Berg *et al.*, 2006)

The large R1 subunit of the class I RNR complex contains two allosteric regulation sites. One of these sites binds only ATP or dATP and regulates activity. Binding dATP inhibits the activity of the RNR complex and binding ATP activates it. The other site is the specificity site, it can bind any of dATP, ATP, dTTP, dCTP and dGTP and the substrate bound will affect which substrates the RNR complex preferentially uses (Larsson et al., 2004). Binding of dATP (or ATP) promotes conversion of UDP and CDP. Binding of dTTP and dCTP promotes the conversion of GDP and inhibits the conversion of UDP and CDP. Binding of dGTP promotes the reduction of ADP.

RNR activity in cells appears to be exquisitely regulated. In addition to allosteric control of activity it has been shown that the transcript level of the R1 subunit is regulated in a cell cycle dependent manner, with the highest transcript level occurring around G1/S phase (Björklund *et al.*, 1990; Gordon and Fantes, 1986).

Class Ia RNR complexes may exist in more than one stoichiometry and these stoichiometries may also influence activity. The R1 subunit has been shown to form a dimer in solution in response to the specificity site being occupied (Ingemarson and Thelander, 1996) and a hexamer if both the specificity and the activity sites are occupied. There have been reports that the R1 hexamer binds a dimer of R2 ($\alpha_6\beta_2$) and is active (Rofougaran *et al.*, 2006) and that the hexameric form is inactive (Fairman *et al.*, 2011). These contradictory results could suggest that the $\alpha_6\beta_2$ form of the complex has two different conformations, one of which is active and one of which is inactive (for more detail see Chapter 4: 4.1.2.1).

Furthermore in yeasts (both *S. pombe* and *S. cerevisiae*) an additional level of regulation is provided by small molecules, which have been reported to bind to the R1 subunit and inhibit RNR activity. In *S. cerevisiae* a factor found to regulate RNR activity was a small, intrinsically disordered protein called Sml1. Sml1 was first identified by Zhao *et al.*, (1998) when they found a mutant strain which was able to grow even in the absence of the essential gene *MEC1*. Sml1 binds to the R1 subunit of the RNR complex (Zhao *et al.*, 1998). In Δ*SML*1 strains the dNTP pool levels were elevated and this effect was not due to an increase in transcription of the RNR proteins. It was subsequently shown by Chabes *et al.*, (1999) that, *in vitro*, purified Sml1 inhibits RNR activity, as measured by a CDP reduction assay and a BIAcore binding experiment using R1 titrated against immobilised Sml1 (Chabes *et al.*, 1999). They then used sucrose gradient centrifugation to test whether Sml1 binding would affect the stoichiometry of the R1 protein. As controls they examined

R1 + dTTP which gave a mixture of dimeric and multimeric R1, and R1 alone, which gave mainly monomeric R1. Adding Sml1 did not shift the distribution of oligomers seen.

In addition to Sml1 other intrinsically disordered proteins have been shown to be involved in RNR regulation: Dif1 is involved in the nuclear import of the small RNR subunit and is degraded in response to DNA damage (Lee *et al.*, 2008). Hug1 binds the small subunit and has been shown to inhibit RNR and accumulate under conditions of high RNR induction (Meurisse *et al.*, 2014).

In *S. pombe* a similar peptide called Spd1 was identified (Woollard *et al.*, 1996). This protein has been shown to bind to both $Cdc22^{R1}$ and $Suc22^{R2}$ and also to inhibit RNR activity (Hakansson *et al.*, 2006; Liu *et al.*, 2003). To investigate whether this inhibition was linked to stoichiometric regulation Nestoras *et al.*, used förster resonance energy transfer (FRET) between $Cdc22^{R1}$ and $Suc22^{R2}$ and showed that in wild type cells a signal could be seen and this was lost in an *spd1* delete strain (Nestoras *et al.*, 2010). This suggests Spd1 has a role in the RNR complex architecture, and since RNR is active in *spd1* delete cells it cannot be that the FRET signal represents an active hetero-tetramer. To establish how *spd1* deletion affected RNR stoichiometry Nestoras *et al*, (2010) used a cross-linked *S. pombe* protein extract which was run through a gel filtration column to see whether two discretely sized populations could be seen. The peak gel filtration fractions, visualised on a western blot, covered a broad range of possible complex sizes suggesting a continuum rather than discrete populations. This may possibly be an artefact of the crosslinking or it may suggest that the complex actually exists in a continuum of oligomeric states rather than discrete populations. They also compared a wild-type extract to an extract from an *spd1* delete mutant and showed that the tail of the fraction peak was slightly shifted towards the higher weights in the *spd1* delete strain. However in hydroxyurea (HU) treated cells the presence or absence of Spd1 had no effect on the weight distribution seen.

A second intrinsically disordered protein, called Spd2, has also been identified in *S. pombe* and shown to also be required for $Cdc22^{R1}$-$Suc22^{R2}$ FRET but not to be required for $Suc22^{R2}$ nuclear import.

Outside of S phase in *S. pombe* $Cdc22^{R1}$ is typically found throughout the cytoplasm of the cell whereas $Suc22^{R2}$ is localised to the nucleus (Liu *et al.*, 2003). During S phase, at least some of the $Suc22^{R2}$ subunit appears to exit the nucleus; leading to the suggestion that one level of RNR regulation is at the level of subcellular compartmentalisation (Liu *et al.*, 2003). Interestingly, $Cdc22^{R1}$ is very highly expressed. A proteomic study by Carpy *et al.*, (2014) suggested that expression peaked during S phase with an average of ~88,000 molecules per cell, reducing to ~50,000 in G2 phase then going back up to ~55,000 in M-phase and ~65,000 in G1 phase. $Suc22^{R2}$ was suggested to be present at lower levels of ~20,000-22,000 molecules per cell on average with the highest levels measured in G2 phase (Carpy *et al.*, 2014). Studies with protein extraction, cross linking and gel filtration have suggested that a hexamer of $Cdc22^{R1}$ might form in *S. pombe* (Nestoras *et al.*, 2010) but the complex was found to be unstable outside of the cell meaning that to truly elucidate the stoichiometries of the RNR complex it would be best to take an *in vivo* measurement. The high density of expression of $Cdc22^{R1}$ combined with the question of the stoichiometry of the RNR complex and the potential that the Spd1 protein may affect this stoichiometry (Nestoras *et al.*, 2010) make it an interesting test case for quantitative PALM.

1.3.3 DNA damage and genomic instability – sensing and responding

One of the defining characteristics of life is the ability to self-replicate, whether that be on a single or multi-cellular level. In order to achieve this the DNA content of cells needs to be copied. This is an inherently inaccurate process and inaccurate replication of the DNA

gives rise to mutations. From a standpoint of evolution, this provides genetic diversity, which is necessary for survival in a changeable world. However, whilst a moderate rate of genetic mutations can be beneficial on a species-wide scale, they can be dangerous on an individual scale, particularly when they become more frequent. In addition to the intrinsic mutations that occur during DNA replication, there are many environmental factors, such as ultra violet (UV) radiation, which can also lead to DNA damage and mutations. Many of these environmental, or extrinsic, factors interact with the DNA replication process and make it more inaccurate, thus increasing the frequency of mutations occurring during DNA replication. An increase in DNA damage and the consequent accumulation of mutations can lead to cell death or the formation of cancers, so it is vitally important that the cell have mechanisms both for sensing and repairing damage.

1.3.3.1 Fork stalling and collapse

One source of genomic instability that can lead to DNA damage is replication fork stalling. The replication fork is the point at which double stranded DNA is being unwound by a helicase and replicated by polymerases. The combination of polymerases, helicases and assorted structural proteins is known as a replication complex or replisome. The fork structure has two arms of single stranded DNA (ssDNA) which act as templates for replication. DNA replication is always in the 5' to 3' direction. The ssDNA strand which is replicated in the same direction as fork progression is termed the leading strand and is replicated continuously by a polymerase. The other ssDNA strand is termed the lagging strand and is replicated discontinuously. A primase "reads" the template strand and creates complementary RNA primers, a DNA polymerase then extends the primed segments. The RNA primers can then be removed and the DNA fragments joined together by DNA ligase.

The progress of the replication fork can be blocked in many ways and these blocks are collectively known as replication fork barriers (RFBs) (reviewed in Lambert and Carr, 2013). RFBs include damaged DNA bases (e.g. by an alkylating agent such as methyl methanesulfonate (MMS)), specific chromosomal features (e.g. centromeres), transcriptional units (e.g. ribosomal DNA), DNA that forms non-canonical structures and poorly characterised regions such as common fragile sites. Certain chemicals can also block replication fork progression, for example HU, which depletes the dNTP pool by inhibiting the RNR complex. Polar RFBs are used by the cells to prevent collisions between the replication and transcription machineries by promoting unidirectional replication, as is the case with ribosomal RFBs (Krings and Bastia, 2004), which are common to all eukaryotes. In *S. pombe* a species specific RFB known as replication termination sequence 1 (RTS1) has evolved to promote unidirectional replication of the mating type locus (Dalgaard and Klar, 2001) and is often used experimentally to study fork stalling.

Replications forks that have been arrested at RBFs but are able to resume replication once the block has been resolved are called stalled forks, forks that are unable to resume replication are known as collapsed forks. Fork collapse can lead to incomplete replication of the DNA, which can be lethal. Cells have three main strategies for coping with RFBs (reviewed in Lambert and Carr, 2013). The first is to prevent the replication fork from arresting at all. To do this there are mechanisms to minimise collisions between replication and transcription machinery (Krings and Bastia, 2004) and to repair DNA damage before replication occurs (Gerald *et al.*, 2002) as well as mechanisms to bypass damage (reviewed in Johnson *et al.*, 1999). The second strategy, when the first has failed to prevent fork arrest, is to attempt to stabilise the fork via the intra-S phase checkpoint (reviewed in Branzei and Foiani, 2010). The advantage of this checkpoint stabilisation is that the stalled fork can rapidly resume replication once the block has been removed and if it cannot resume replication the nascent DNA ends are protected until the stalled fork can merge with a

converging fork. Activation of the intra-S phase checkpoint relies on the presence of enough ssDNA to cause RPA binding (Zou and Elledge, 2003), where insufficient ssDNA is produced the fork cannot be stabilised and will collapse leaving the nascent ends unprotected. When this occurs the fork must be rescued in a process which has been shown to require HR proteins (Costanzo, 2011; Petermann *et al.*, 2010). However rescue of collapsed forks by HR has a cost of potentially increasing gross chromosomal rearrangements (Lambert *et al.*, 2005).

1.3.3.2 Double strand breaks

One form of DNA damage is a double strand break (DSB) which is highly toxic and can lead to large scale loss of downstream genetic material if unrepaired (for a review on single strand breaks see (Caldecott, 2008)). DSBs can occur spontaneously as a result of fork stalling or collapse due to ssDNA nicks or inter-strand cross-linking (ICL), they can occur deliberately during meiosis or they can occur as a result of environmental factors such as ionising radiation (reviewed in Mehta and Haber, 2014). Once a DSB has formed there are two main pathways for repair: homologous recombination (HR), in which the 3' tail of one of the strands invades a homologous sequence which it uses as a template to continue replication (reviewed by San Filippo *et al.*, 2008), and non-homologous end joining (NHEJ) in which the ends either side of the break are ligated, this method can lead to errors such as insertions or deletions (reviewed by Lieber, 2010).

1.3.3.3 Cell cycle checkpoints

To ensure that DNA replication has proceeded fully, and that DNA damage has been repaired, before the cell enters mitosis, checkpoint mechanisms have evolved that can arrest

cell cycle progression at various points in the cell cycle. The cell cycle can be arrested by checkpoint pathways at the G1/S boundary – this allows time for DNA repair before replication is initiated. Arrest can also be imposed at the G2/M boundary, which allows DNA repair to complete before chromosome segregation. In addition, progression through S phase can be inhibited by the intra-S checkpoint pathway (reviewed in Houtgraaf *et al.*, 2006). This allows cells to coordinate repair events with progress through S phase.

In *S. pombe* these three DNA structure-dependent checkpoints use a common DNA damage/DNA structure sensing mechanism. A transition between ds and ssDNA is sensed by the loading of the Rad9-Hus1-Rad1 (911) complex onto ssDNA by the Rad17-RFC3 complex (Ellison and Stillman, 2003) (Figure 1.10). This process is dependent on the presence of replication protein A (RPA) (Zou and Elledge, 2003), a protein with a high affinity for ssDNA. Stretches of ssDNA can be created by nuclear resection at DSBs or by the helicase unwinding of dsDNA becoming decoupled from replication at a stalled replication fork. RPA then binds to these stretches of ssDNA. Rad17/RFC(2-5) then recognises this RPA/ssDNA complex and loads the 9-1-1 complex (Caspari and Carr, 1999). Rad9, a subunit of the 9-1-1 complex, then recruits TopBP1 (also known as Rad4 and Cut5 in *S. pombe* and as Dpb11 in *S. cerevisiae*). Rad3 (*S. cerevisiae* Mec1, Human ATR) is a phosphatidylinositol 3-kinase like kinase (PIKK) (Bentley *et al.*, 1996) that forms an obligate dimer with Rad26. Rad26 is another RPA binding protein that recruits Rad3 to ssDNA. Once both TopBP1 (Rad9-dependent) and Rad3-Rad26 are co-localised to the same stretch of ssDNA, TopBP1 can bind to and activate Rad3 (Delacroix *et al.*, 2007). Once activated in this way, Rad3 is able to phosphorylate a wide range of substrates (Willis *et al.*, 2016), including both Chk1 (Walworth and Bernards, 1996) and Cds1 (Lindsay *et al.*, 1998), two effector kinases. The effector kinases are recruited to the checkpoint by mediator proteins Crb2 (*S. cerevisiae* Rad9, human 55BP1) and Mrc1.

The ultimate action of these checkpoints is to inactivate the Cdc2 (human CDK1)-cyclin-B kinase complex. When active this complex drives cellular progression into mitosis (Gould and Nurse, 1989). Progress into mitosis is inhibited when Cdc2 is phosphorylated on the tyrosine 15 and threonine 14 residues (Norbury *et al.*, 1991). Cdc2 activity is also regulated through the protein Cdc25. Cdc25 is an activator of Cdc2 which works by de-phosphorylating it (Gautier *et al.*, 1991). Inhibition of Cdc25 is achieved, in part, by phosphorylation, which induces Cdc25 to interact with 14-3-3 proteins and causes it to be excluded from the nucleus (Kumagai and Dunphy, 1999; Lopez-Girona *et al.*, 1999; Zeng and Piwnica-Worms, 1999).

In the G2/M DNA damage checkpoint Cdc2 inhibition is achieved mainly through the Chk1 effector kinase (Martinho *et al.*, 1998). Rad3 phosphorylates Crb2 which recruits Chk1 to be activated by Rad3. Chk1 then activates both Mik1 and Wee1 (Boddy *et al.*, 1998; O'Connell *et al.*, 1997). Mik1 and Wee1 inactivate Cdc2 by phosphorylation of the Tyr-15 residue (O'Connell *et al.*, 1997). At the same time Chk1 phosphorylates Cdc25 causing it to bind to 14-3-3 and be excluded from the nucleus. Chk1 also induces degradation of the small regulatory protein Spd1. Spd1 is an inhibitor of the ribonucleotide reductase (RNR) (see 1.3.2), in its absence RNR is active and can produce dNTPs which are needed for DNA repair.

In the intra-S phase checkpoint Cdc2 inhibition is achieved mainly through the Cds1 effector kinase. (Christensen *et al.*, 2000; Lindsay *et al.*, 1998). Mrc1 recruits Cds1 to be activated by Rad3. Activation of Cds1 leads to Mik1 accumulation (Christensen *et al.*, 2000). Mik1 phosphorylates Cdc2 (as in the G2/M checkpoint). There is also evidence that Cds1 phosphorylates Cdc25 contributing to its nuclear exclusion and is able to directly inhibit Cdc25 (Furnari *et al.*, 1999). Cds1/Mrc1 also acts to stabilise stalled forks (Lopes *et al.*, 2001), inhibit the firing of origins (Kumar and Huberman, 2009) and regulate the

Mus81-Eme1 endonuclease (Kai *et al*., 2005).

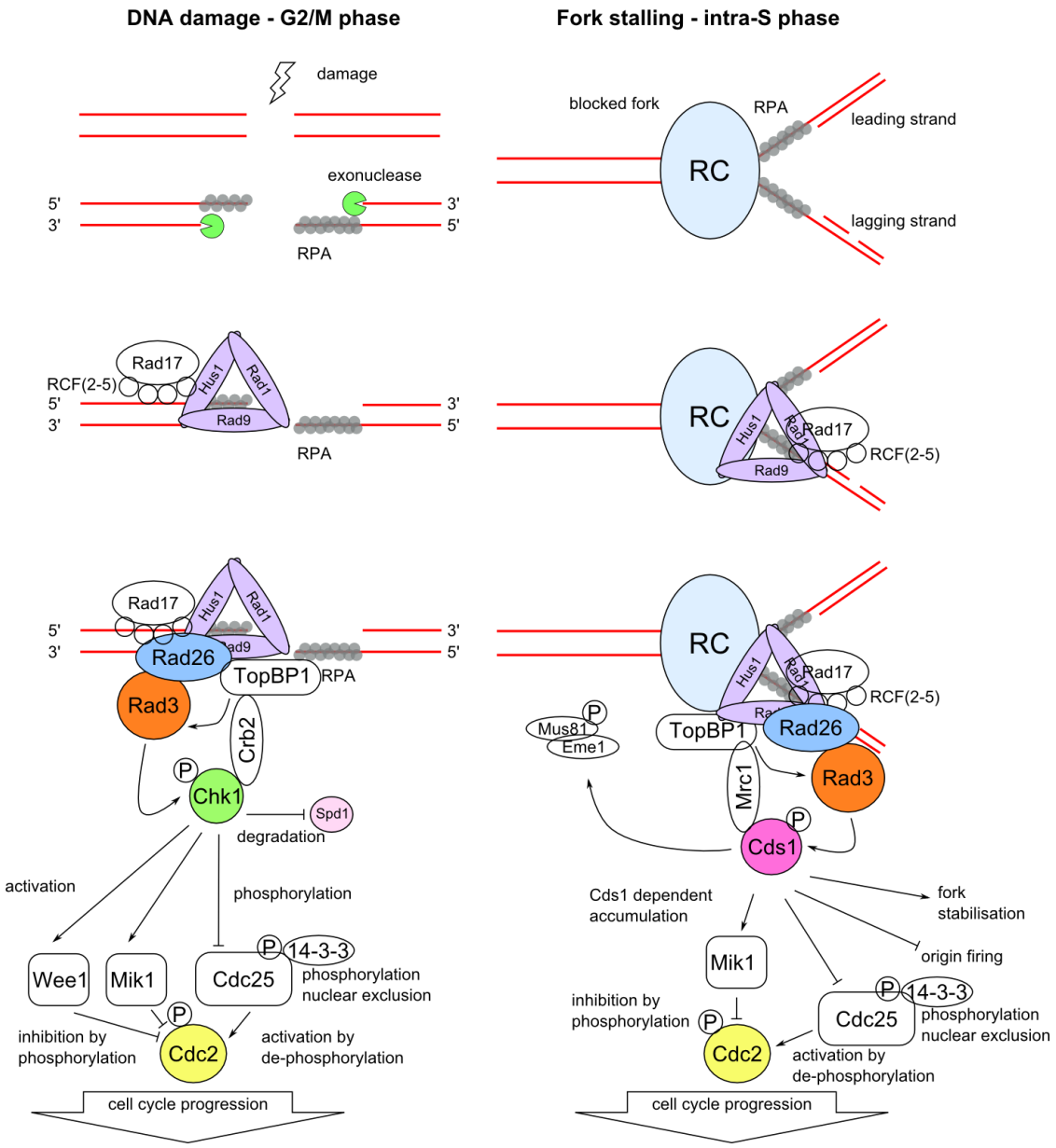
For a concise review on DNA checkpoints in fission yeast see Furuya and Carr, (2003).

Figure 1.10 Simplified schematic for the G2/M and intra-S phase checkpoints in *S. pombe*

DNA DSBs and fork stalling can both lead to the exposure of ssDNA. This ssDNA is immediately coated with RPA (grey circles). In response to an increase in RPA, Rad17 in complex with four RFC subunits (2-5) recruits the 9-1-1 complex to the junction between ss and ds DNA. The 9-1-1 complex is a PCNA-like sliding clamp formed of Rad9, Rad1 and Hus1. Independently. Rad9 recruits another protein, TopBP1. The Rad3-Rad26 complex is also recruited to ssDNA coated with RPA and TopBP1 activates Rad3.

In the G2/M checkpoint Rad3 is then able phosphorylate Chk1. Chk1 can activate Wee1 and Mik1. Wee1 and Mik1 phosphorylate Cdc2 on the Y15 residue, deactivating it and preventing progression into mitosis. Chk1 also phosphorylates Cdc25, causing it to be excluded from the nucleus and preventing it from de-phosphorylating (and therefore activating) Cdc2. Additionally Chk1 induces the degradation of Spd1 which is an inhibitor of the RNR complex.

In the Intra-S phase checkpoint Rad3 phosphorylates Cds1. Cds1 causes accumulation of Mik1, which inhibits Cdc2 by phosphorylation on the Y15 residue. Cds1 also acts to inhibit origin firing, to stabilise the replication fork and to regulate Mus81/Eme1.


DNA damage - G2/M phase
damage
exonuclease
5'
3'
3'
5'
RPA
Rad17
RCF(2-5)
Hus1
Rad1
Rad9
Rad26
TopBP1
Rad3
Crb2
Chk1
Spd1
degradation
activation
phosphorylation
Wee1
Mik1
Cdc25
14-3-3
phosphorylation
nuclear exclusion
inhibition by
phosphorylation
Cdc2
activation by
de-phosphorylation
cell cycle progression
Fork stalling - intra-S phase
blocked fork
RPA
RC
leading strand
lagging strand
Mus81
Eme1
Mrc1
Cds1
Cds1 dependent
accumulation
fork
stabilisation
origin firing
Mik1
inhibition by
phosphorylation
Cdc25
phosphorylation
nuclear exclusion
Cdc2
activation by
de-phosphorylation
cell cycle progression

1.3.3.4 HR repair of double strand breaks

Once the checkpoint has been activated as a result of damage the cell must then activate a damage repair response. Checkpoint activation and the repair response have been suggested to be linked (reviewed in Kara A. Nyberg *et al.*, 2002; Zhou and Elledge, 2000). As mentioned above there are two major repair pathways for DNA DSBs: HR and NHEJ. HR is an important biological process not only for repairing DNA damage (Figure 1.11), in eukaryotic organisms it also plays a part in generating genetic diversity by allowing the exchange of genetic material between paternal and maternal chromosomes during meiosis. HR is the more accurate of the two pathways for repairing double strand breaks as it involves one end of the break invading a homologous sequence and using it as a template for replication (reviewed in Thompson and Schild, 2001).

When a DSB is formed the first step in repair is resection of the ends, in which an endonuclease removes DNA bases starting from the 5' end of the break leaving a 3' overhang of ssDNA. This ssDNA is immediately coated by RPA which binds to it with a high affinity (reviewed in Wold, 1997). In order for HR to occur, Rad51 replaces RPA as the main protein that binds to the ssDNA and forms a nucleoprotein filament (Sauvageau *et al.*, 2005). It is this Rad51 filament that promotes strand invasion strand invasion of ssDNA in the duplex DNA of a sister chromatid or other homologous sequence (Sung and Robberson, 1995). The Rad52 protein targets Rad51 to RPA and aids in filament formation of Rad51 (New *et al.*, 1998). In addition to Rad52, the formation of the Rad51 nucleoprotein filament requires various mediator proteins; the Rad55/Rad57 heterodimer complex (homologous to S. *cerevisiae* Rad55/57) which stabilises the Rad51 nucleoprotein filament to enhance strand exchange (Sung, 1997); and, in *S. pombe*, the Swi5/Sfr1 complex which has the same function as Rad55/57 and acts in parallel with but independently of it (Akamatsu *et al.,* 2007). Once the nucleoprotein filament is formed the filament is able to

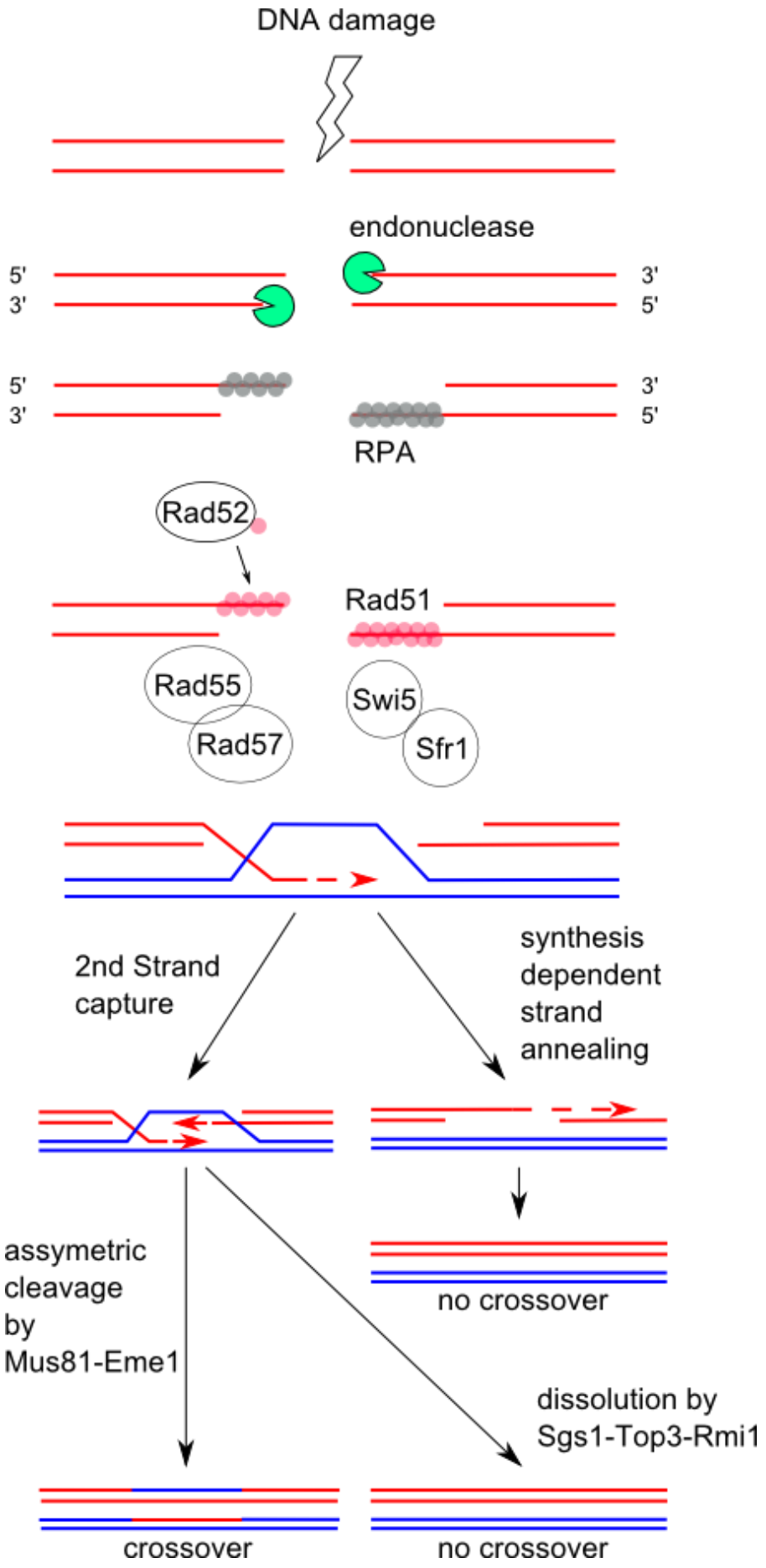
invade a sister chromatid or other homologous sequence to use as a template for replication (Sung and Robberson, 1995). After the first strand invasion there are two models for what is thought to happen: One model is that there is second end capture leading to the formation of a double Holliday junction structure that is either resolved by cleavage by Mus81-Eme1, leading to crossover products, or by dissolution by Sgs1-Top3-Rmi1. The other model is that there is a continuously migrating D-loop and that eventually the invading strand is displaced and anneals with the second resected DSB end, giving no crossover products (reviewed by San Filippo *et al.*, 2008).

Recently two proteins with no previously described function were identified in a search for homology with the *S. cerevisiae* protein Ris1, which is a DNA-dependent ATPase with roles in DNA repair and gene silencing (Dziadkowiec *et al.*, 2009). These two proteins, Rrp1 and Rrp2, are involved in the Swi5/Sfr1 arm of Rad51-dependent HR and have been shown to form foci (clusters of multiple, co-localised molecules) at sites of DNA damage, when overexpressed from a plasmid (Dziadkowiec et al., 2009, 2013). These proteins have a very low abundance, making it difficult to observe them at endogenous levels using conventional fluorescence microscopy. This low abundance and potential DNA binding activity of Rrp1 and Rrp2 proteins makes them an interesting target for PALM imaging to discover whether improved detection and resolution would enable imaging of Rrp1 or Rrp2 at endogenous levels and allow further elucidation of their function in HR.

Figure 1.11 Homologous recombination repair of double strand breaks

Adapted from (San Filippo *et al.*, 2008)

When DNA damage causes a double strand break exonucleases resect the 5' ends to leave 3' overhangs of ssDNA. This ssDNA is coated with RPA which acts to stabilise the ssDNA. In order for HR to occur the RPA needs to be replaced with Rad51, however the RPA binds with a higher affinity than the Rad51. To overcome this Rad51 requires Rad52 and either or Rad55/57 or Swi5/Sfr1 in order to bind to the ssDNA and form a nucleoprotein filament. This nucleoprotein filament is capable of performing strand invasion into a homologous sequence, usually on the sister chromatid, which can be used as a template for replication across the broken section. After this strand invasion there are a number of different routes to finishing the repair with varying outcomes. One possibility is synthesis dependent strand annealing, in which there is no second end capture and after extension of the invading strand by a DNA polymerase it is displaced and returns to anneal with the 3' overhang of its original partner strand. This leads to non-crossover products. Alternatively the second ssDNA end can be captured by the homologous sequence forming a double Holliday junction. This can be resolved through asymmetric cleavage by Mus81-Eme1 leading to crossover products or through dissolution by Sgs1-Top3-Rmi1 leading to non-crossover products.


DNA damage
endonuclease
5'
3'
3'
5'
5'
3'
3'
5'
RPA
Rad52
Rad51
Rad55
Rad57
Swi5
Sfr1
2nd Strand
capture
synthesis
dependent
strand
annealing
assymetric
cleavage
by
Mus81-Eme1
no crossover
dissolution by
Sgs1-Top3-Rmi1
crossover
no crossover

2. Materials and Methods

2.1 Growth media

All media were filter sterilised using a Steritop™ 0.22 µm 1 L filter unit (Merck Millipore)

2.1.1 Yeast growth media

2.1.1.1 Yeast extract (YE) Media

Component	1 Litre
Yeast Extract	5 g
Glucose	30 g
Adenine	0.8 g
Leucine	0.4 g
Uracil	0.4 g
Histidine	0.4 g
Arginine	0.4 g

2.1.1.2 Edinburgh minimal media (EMM)

Component	1 Litre
20x EMM Salts	50 ml
20% NH_4Cl	25 ml
0.4 M $NaHPO_4$	25 ml
40% Glucose	50 ml
1000x Vitamins	1 ml
10,000x Trace Elements	100 µl

20x EMM Salts

Component	1 Litre
$C_8H_5KO_4$	61.2 g
KCl	20 g
$MgCl_2.6H_2O$	21.4 g
Na_2SO_4	0.2 g
$CaCl_2.2H_2O$	0.26 g

10,000x Trace Elements

Component	**100 ml**
H_3BO_3	500 mg
$MnSO_4$	400 mg
$ZnSO_4.7H_2O$	400 mg
$FeCl_2.6H_2O$	200 mg
Na_2MoO_4	150 mg
KI	100 mg
$CuSO_4.5H_2O$	40 mg
Citric Acid	1000 mg

1000x Vitamins

Component	**100 ml**
Pantothenic acid	100 mg
Nicotinic Acid	1 g
Inositol	1 g
Biotin	1 mg

2.1.1.3 Other stock supplements

Adenine	100 g/L	(100×)
Leucine	100 g/L	(100×)
Uracil	20 g/L	(20×)

2.1.1.4 Agar plates

EMM agar plates were made up by autoclaving 10 g Bacto Agar and 300 ml double distilled H_2O (ddH_2O) then adding 100 ml of 4× EMM stock + relevant supplements to the molten agarose mixture and pouring into sterile petri dishes. YE agar plates were made up in the same way but using 100 ml of a 4× YEA stock instead.

2.1.2 Bacterial growth media

2.1.2.1 Lysogeny broth (LB) media

Component	1 L
Tryptone	10 g
Yeast Extract	5 g
Sodium Chloride	5 g

2.1.2.2 Agar plates

Agar plates were made up of 1 L LB + 12 g Difco Bacto Agar

2.1.3 Drugs used for genetic selection

Drug	Final Concentration
Nourseothricin sulphate (NAT)	100 μg/ml
Geneticin dislulphate (G418)	200 μg/ml
5-fluoroorotic acid (5-FOA)	1 mg/ml
Ampicillin sodium salt (AMP)	100 μg/ml

2.1.4 Genotoxic agents

Hydroxyurea (HU)

Methyl methanesulfonate (MMS)

Camptothecin (CPT)

2.2 General molecular cloning techniques

2.2.1 PCR amplification of DNA fragments

To amplify DNA fragments the KOD Hot-start polymerase (Novagen) was used as follows:

Component	**Volume for 50 µl reaction**
10x KOD buffer	5 µl
dNTPs (2 mM)	5 µl
25 mM MgSO4	3 µl
KOD Hot-start polymerase	1 µl
DNA (plasmid ~10 ng/µl, genomic ~50-100 ng/µl)	1 µl
Forward and Reverse primers (20 µM)	0.5 µl
ddH_2O	Make reaction up to 50 µl

For amplifications from genomic DNA for which KOD was unable to produce a product TaKaRa Ex Taq (TaKaRa Clontech) was used:

Component	**Volume for 50 µl reaction**
10x buffer	5 µl
dNTPs (2 mM)	4 µl
Ex Taq polymerase	0.25 µl
DNA (plasmid ~10 ng/µl, genomic ~50-100 ng/µl)	1 µl
Forward and Reverse primers (20 µM)	1 µl
ddH_2O	Make reaction up to 50 µl

The thermocycling regime used was:

1	95 °C	180 s
2	95 °C	15 s
3	57 °C	30 s
4	72 °C	90 s

Stages 2-4 were repeated 29x after initial cycle

PCR products were run on a 1% (w/v) agarose gel containing 0.004% EtBr for visualisation against GeneRuler™ DNA ladder Mix (ThermoFisher Scientific). PCR products for gel extraction were also run on a 1% (w/v) agarose gel containing 0.004% EtBr and the products extracted using a Nucleospin® gel and PCR clean-up kit (Machery Nagel) according to the manufacturer's instructions. PCR products that were not for gel extraction were also cleaned up using the Nucleospin® gel and PCR clean-up kit (Machery

Nagel).

2.2.2 Restriction digest

Restriction enzymes were all purchased from New England Biolabs. Restriction digests were performed in the recommended buffer at the concentrations suggested by the manufacturer. Digests were carried out at 37 °C for 1 hr. Digested fragments were purified using a Nuscleospin® Gel and PCR Clean-up kit (Machery Nagel) or by running on a 1% agarose gel and gel extracting and purifying using the same kit.

2.2.3 DNA ligation

Ligations of restriction digested PCR products into a plasmid vector were performed using T4 DNA ligase (NEB). Ligation reaction mixtures were prepared as per the manufacturer's instructions and incubated overnight at 16 °C. An insert to vector ratio of 1:1 was used.

2.3 General *E. coli* techniques

2.3.1 Transformation

To transform ligation products into *E. coli* cells 50 µl of competent DL21 cells were added to the ligation mixture and incubated on ice for 10 minutes before being heat-shocked at 42 °C for 60 seconds then returned to ice for 5 minutes. 450 µl of LB were then added to the cells and the mixture incubated at 37 °C for 1 hour then 100 µl of cell mixture were plated onto LB agar (LBA) plates containing the relevant selective agent at the concentrations detailed in 2.1.3.

2.3.2 Extraction of plasmid DNA

Single *E. coli* colonies were picked from the selective media plates and grown overnight at 37 °C in 5 ml LB media with relevant antibiotic. Stock patches were created on fresh antibiotic LBA plates from the liquid culture.

Small scale plasmid DNA extractions were performed either using a QIAprep Spin Miniprep Kit (Qiagen) following the manufacturer's instructions or by the following protocol: Cultures were spun at 3600 × g for 10 minutes, the supernatant was discarded and the cells re-suspended in 200 µl of P1 (50 mM TrisCl pH8, 10 mM EDTA, 100 µg/ml RNAse A) and transferred to a 1.5 ml tube. 300 µl of P2 (200 mM NaOH, 1% SDS w/v) were added and incubated for 5 minutes at room temperature then 300 µl of P3 (3 M KOAc pH5.5) were added and the mixture incubated 5 minutes on ice. The mixture was then spun at 22,000 × g for 10 minutes and the supernatant transferred to a fresh 1.5 ml tube and 700 µl of isopropanol were added, then the mixture was incubated at -20 °C for 10 minutes and then spun again at 22,000 × g for 10 minutes. The supernatant was tipped off and the resulting pellet washed in 700 µl of 75% (v/v) ethanol before being dried in a vacuum centrifuge for 10-15 minutes. Pellets were re-suspended in 19 µl ddH2O + 1 µl RNAse A (from stock 10 mg/ml).

Large scale plasmid DNA extractions were performed using a Nucleobond® Xtra Midi kit (Machery Nagel) according to the manufacturer's instructions.

2.4 General *S. pombe* cell biology techniques

2.4.1 Crossing strains

Crossing was performed in order to combine genes encoding fluorescently tagged fusion proteins with other genetic backgrounds. *S. pombe* strains of opposite mating type were freshly grown on YEA plates before a loopful of cells from each strain were taken and mixed on ELN plates with 10 µl of ddH_2O and then incubated at 25 °C for 2-3 days. Asci formation was checked under a widefield, white light microscope and, where asci were present, a loopful of cells were re-suspended in 1 ml ddH_2O with 2 µl Glusulase (Perkin Elmer) in order to digest the cell walls, freeing asci and destroying un-mated cells. Spores were counted on a haemocytometer and then 500 were plated on YEA and incubated at 30 °C for 2-3 days.

Strains containing all genes of interest were selected by replica-plating onto YEA containing relevant selective media and by DNA extraction and PCR to screen for the presence of a fluorescent label by size difference compared to the wild type.

Where the mating types of two strains to be crossed were the same one strain was back-crossed to either AW309 (h-) or AW310 (h+) (see Appendix Table 8-1 for genotypes). Mating type was checked by crossing selected colonies with both AW309 and AW310 and checking for asci formation.

2.4.2 LiAcTE Transformation

Cells for transformation were cultured overnight in 10 ml YE to a density of $1x10^7$ cells/ml at 30 °C. 10 ml of culture per transformation reaction were spun down at 3,200 × g for 5 minutes then washed in 5 ml ddH_2O and spun down for another 5 minutes. The cell pellet was then re-suspended in 1 ml ddH_20 and transferred to a 1.5 ml tube and spun at

13,000 × g for 1 minute. The pellet was washed in 1 ml filter sterilised (0.22 μm filter) LiAc-TE (0.1 M litium acetate, 10 mM Tris pH 7.5, 1 mM EDTA), spun at 13,000 × g for 1 minute then re-suspended in 100 μl LiAc-TE. 2 μl of carrier DNA (salmon sperm DNA, 10 mg/ml) and up to 10 μl of DNA (plasmid or PCR product) were added and the mixture incubated for 10 minutes on the bench. 40% (w/v) polyethylene glycol (PEG) was filter sterilised (0.22 μm filter) and 260 μl added to the sample which was gently vortexed then incubated at 30 °C for 60 minutes. 43 μl of dimethyl sulfoxide (DMSO) were added and the mixture heat-shocked at 42 °C for 5 minutes then spun down at 10,000 × g for 1 minute, washed once in 1 ml ddH_2O, re-suspended in 200 μl ddH_2O then 100 μl plated onto either YEA or EMM containing the relevant amino acid supplements.

2.4.3 Base strain creation for recombination mediated cassette exchange

The gene replacement base strain for recombination mediated cassette exchange was created by designing primers to amplify the lox sites and *ura4* marker gene from the plasmid pAW1 and to have ~80 bp homology with the area around the gene of interest. The resulting linear fragment was run in a 1% agarose gel for extraction and purified using a Nucleospin® Gel and PCR Clean-up kit (Machery Nagel). The fragment was transformed into the uracil (*ura4* D-18) and leucine (*leu* 1-32) deficient strain AW310 (see strain list in Appendix) using the LiAcTE transformation protocol and colonies were selected for on EMM plates containing leucine but not uracil. Successful colonies were screened using PCR with primers to the *ura4* marker gene and a region outside of the site of integration then sequenced to check the marker sequence had gone into the correct locus.

2.4.4 Recombination mediated cassette exchange

Recombination mediated cassette exchange was performed as described in (Watson *et al.*, 2008). The gene of interest was amplified by PCR, using primers to add restriction sites for cloning and sub-cloning, and cloned into the vector plasmid pAW8 (which contains the *leu1* marker gene and the Cre-recombinase gene under the control of the *nmt1* thiamine repressible promoter), between the flanking lox sites. The sequence encoding the desired fluorescent protein tag was then sub-cloned into the plasmid, such that it was in-frame with the gene sequence, either at the N or the C terminal of the gene of interest. The resulting plasmid was then transformed into base strain cells using the LiAcTE transformation protocol. Successful transformants were selected for on EMM plates with no supplemental amino acids. These cells were then grown overnight at 30 °C in 10 ml YE media (which is sufficient to allow expression of the Cre-recombinase from the pAW8 plasmid) and then counted on a haemocytometer. 10,000 cells were plated onto YEA containing 5-FOA. Cells which successfully grew on 5-FOA (indicating loss of the *ura4* marker gene) were patched out and assessed for the presence of the tag by DNA extraction and PCR.

2.4.5 Lactose gradient enrichment for S phase cells

A solution of 20% lactose in phosphate buffered saline (PBS) was prepared and 10 ml aliquots in 15 ml falcon tubes were frozen upright at -20 °C overnight. One aliquot per culture was defrosted upright at room temperature for ~30 minutes, this allowed the formation of a gradient of lactose concentrations. Cells were grown in EMM supplemented with adenine and uracil to a density of 0.5-1 × 10^7 cells/ml as measured using a haemocytometer. 100 ml of cells were pelleted by centrifugation at 2600 × g (3000

rpm Beckman Coulter, Allegra x-22 series) for 3 minutes and the pellet was re-suspended in 400 µl of PBS. Using a truncated pipette tip the cells were layered on top of the lactose gradient which was then centrifuged at 954 × g (1100 rpm Beckman Coulter, Allegra x-22 series) for 10 minutes, forming a band of cells part way down the gradient. 400 µl of cells were collected from the top of this band, 500 µl of PBS were added to the cells and the cells were pelleted in a benchtop centrifuge then re-suspended in 1 ml YE media. 10 µl of cells were added to a microscope slide to check the size and synchrony of the cells, with the aim being to have a culture of predominantly small G2 cells. For experiments involving DNA damage induction the cells were split into two tubes and made up to 1 ml with YE, one tube was treated with MMS or HU and the other tube left untreated. Cells were then incubated at 30 °C with shaking for 2.5 hrs until septation was observed in the majority of cells, indicating S phase. Cells were spun down in benchtop centrifuge and re-suspended in PBS. At this stage if two populations were to be visualised simultaneously one was treated with a ConA Alexa Fluor® 647 conjugate dye (see 2.5.5) before being washed three times in PBS. Cells were then treated with ice cold sodium azide (a mitochondrial inhibitor) at a concentration of 0.01% to prevent active removal of proteins from the DNA and washed a further three times with PBS. If multiple populations were being visualised simultaneously they were mixed together in a single tube. Cells were spun down and re-suspended in a small volume of PBS and kept on ice. A thin pad of 1% agarose was prepared by sandwiching 100 µl molten agarose between two 18 × 18 mm cover slips. 5 µl of cells were pipetted onto one side of this agarose pad which was then placed face down onto a 40 × 20 mm glass slide.

2.4.6 Strain viability assay

Cells were grown overnight in liquid YE media to mid log phase. They were then diluted

and grown for at least another cell cycle to give an OD_{595} of ~1 (roughly 1 x 10^7 cells/ml). In a 98 well plate each cell culture was diluted to give an OD_{595} of either 1 or the lowest value of the cultures (i.e.; if one culture had only grown to 0.8 and the rest were at 1 they would all be diluted to 0.8). Four serial dilutions of 1/10 were then performed in water to give a total of five different dilutions. 7 µl of each dilution were spotted onto YEA plates containing varying concentrations of DNA damaging agents. The agarose plates were composed of 50 ml YEA mixed with the relevant DNA damaging agents: HU was prepared as a 1 M stock from powder (Fisher) dissolved in water and stored at -20 °C in 1 ml aliquots. MMS was prepared as a 1% stock from liquid (Sigma).

2.4.7 Extraction of genomic DNA

Cells were cultured overnight in 10 ml YE media and 1 ml of culture centrifuged at 13,000 × g, the supernatant was discarded and the pellet re-suspended in 1 mg/ml lyticase (Fisher) in CSE (citrate phosphate buffer (citric acid 20 mM, dibasic sodium phosphate 50 mM) pH 5.6, Sorbitol 1.2 M, EDTA 40 mM) and then incubated for 30 minutes at 37 °C. Cells were then pelleted and re-suspended in 450 µl of 5×TE buffer and 50 µl 10 % (w/v) SDS and incubated at room temperature for 5 minutes. 150 µl KAc (5M) were added and the sample put on ice for 5 minutes before being spun down at 13,000 × g at 4 °C for 10 minutes. 500 µl of supernatant were transferred to 500 µl isopropanol and put on ice for 10 minutes. The sample was spun at 16,000 × g for 10 minutes at 4 °C and the supernatant was tipped off. The resulting pellet was washed in 500 µl 75% ethanol, spun for 5 minutes at 13,000 × g at 4 °C, the supernatant was tipped off and the pellet was dried in a vacuum centrifuge for 10-15 minutes. The dry pellet was re-suspended in 200 µl ddH_2O.

2.4.8 Whole cell protein extraction

Cells were grown to mid log phase (~1×10^7 cells/ml, OD_{595} = 1) and 5 ODs worth of cells were spun down at 3,200 × g for 5 minutes. The pellet was washed in 1 ml of ddH_2O and transferred to a 1.5 ml tube then spun down at 16,000 × g for 1 minute. The supernatant was discarded and the pellet washed in 200 µl of 20% (w/v) trichloro acetic acid (TCA) and spun down at 16,000 × g for 1 minute. The supernatant was discarded again and a 1.5 ml tube-capful of acid-washed glass beads (Sigma-Aldrich) and 200 µl TCA added to the pellet. Cells were then broken open using a ribolyser. The resulting lysate was transferred to clean tubes by using a heated needle to puncture a hole in the bottom of the tube, placing the punctured tube into another tube and briefly spinning in a centrifuge. The lysate was then spun at 16,000 × g for 5 minutes and all supernatant was removed. The pellet was re-suspended in 1× sample buffer (50 mM Tris pH6.8, 2% SDS, 10% glycerol, 12.5 mM EDTA, 0.02% bromophenol blue, 1% β-mercaptoethanol) and boiled for 5 minutes at 100 °C.

2.4.9 Western blot

Polyacrylamide gels were poured with an 8% resolving gel. Gels were run in 1 × SDS running buffer (0.025 M Tris Base, 0.25 M Glycine, 0.1% SDS) at 80 volts through the stacking gel and 120 volts through the resolving gel.

Protein samples were transferred to a Nitrocellulose membrane (GE Healthcare, Nitrocellulose, Hybond RPN3032D) in 1 × transfer buffer (20 mM Tris Base, 750 mM Glycine, 20% Methanol, 0.025% SDS) for 2 hours at constant 300 mA. Transfer was confirmed by Ponceau-S solution (0.2% (w/v) Ponceau S and 3% (w/v) TCA in ddH_2O) staining of the membrane.

The membrane was then blocked in 5% (w/v) milk (Marvel) in PBS with 0.1% tween (PBST) overnight at 4 °C. Where the antibody for mEos was being used a membrane with extracts from wild type cells was also produced which was blocked in 5% milk PBST for 30 minutes before being incubated overnight at 4 °C with 5% milk PBST containing 1/1000 rabbit anti-mEos in an attempt to reduce the non-specific binding of the antibody. The membrane containing the samples of interest was washed 3 × 10 min in PBST then incubated for 1 hr at room temperature with 5% milk PBST containing 1/1000 of the primary antibody. The membrane was then washed a further 3 × 10 min in PBST before being incubated for 1 hr at room temperature with 5% milk PBST containing 1/5000 of the horse radish peroxidase (HRP) conjugated secondary antibody. The membrane was then washed 3 × 10 min in PBST and treated with ECL Plus Western Lightning (Perkin Elmer) before being exposed to a light sensitive film (GE Healthcare Hyperfilm ECL) which was then developed.

Resolving gel (8%)	**10 ml**
H_20	4.6 ml
30% acrylamide mix	2.7 ml
1.5 M Tris pH8.8	2.5 ml
10% SDS	0.1 ml
10% APS	0.1 ml
TEMED	0.006 ml

Table 2-1 8% resolving gel

Stacking gel	**4 ml**
H_20	2.7 ml
30% acrylamide mix	0.67 ml
1.5 M Tris pH8.8	0.5 ml
10% SDS	0.04 ml
10% APS	0.04 ml
TEMED	0.004 ml

Table 2-2 Stacking gel

Antibody	Type	Manufacturer	Dilution
Anti-mEos	Rabbit polyclonal	Custom made	1/1000
Swine anti-rabbit HRP	Swine polyclonal	DakoCytomation	1/5000
Anti-Cdc22	Goat polyclonal	Custom made	1/1000
Rabbit anti-goat HRP	Rabbit poly-clonal	DakoCytomation	1/5000

Table 2-3 Antibodies used

2.5 *S. pombe* sample preparation for single molecule localisation microscopy and fluorescence correlation spectroscopy

2.5.1 Preparation of slides and an agarose pad

20 × 40 mm glass slides and 18 × 18 mm glass coverslips were cleaned using a combination of ozone and UV for at least 2 hours before use. An agarose pad was used in order to immobilise cells on the slide which was produced by melting 1% (w/v) agarose in ddH_2O or PBS and sandwiching 50 µl of it between two cleaned 18 × 18 mm coverslips to set. Once set one coverslip was prised off the agarose pad and the cells applied to the surface of the pad. The pad was then placed cells side down onto a clean 20 × 40 mm glass slide.

2.5.2 Fixed cells

Cell cultures were grown to an OD_{595} of 0.1-0.5 then 5 ODs worth spun down and washed 3 × in 1ml phosphate buffered saline (PBS). Cells were then incubated at room temperature in 1 ml 1% (v/v) formaldehyde (Thermo Fisher) in PBS for 30 minutes before being washed 3 × in 1 ml PBS. The cell pellet was then re-suspended in 10-50 µl PBS. 10 µl of cells were applied to the agarose pad which was placed cell side down onto a clean slide.

2.5.3 Unfixed cells

Cell cultures were either grown asynchronously to an OD_{595} of 0.1-0.5 as for fixed cells or else synchronised using a lactose gradient (2.2.5) and then washed 3 × in PBS before being re-suspended in 10-50 µl PBS and added to the agarose pad. For motion blur imaging after initial washing the cells were suspended in 1 ml PBS and 10 µl 10% (w/v) sodium azide (a mitochondrial inhibitor) were added (to prevent active unloading of molecules from the DNA) and the cells incubated on ice for 1 minute. The cells were then washed 3 × in 1 ml ice cold PBS and re-suspended in 15 µl PBS. Cells were kept on ice prior to imaging.

2.5.4 Cell lysate

Cells were grown to mid log phase (~1×10^7 cells/ml, OD_{595} = 1) and 5 OD_{595}s worth of cells were spun down at 3,200 × g for 5 minutes. The pellet was washed in 1 ml of ddH_2O and transferred to a 1.5 ml tube then spun down at 16,000 × g for 1 minute, the supernatant was discarded and the pellet washed in 200 µl of PBS + proteinase inhibitor cocktail (Roche) and spun down at 16,000 × g for 1 minute. The supernatant was discarded again and a 1.5 ml tube-capful of acid-washed glass beads (Sigma-Aldrich) and 200 µl PBS + proteinase inhibitor cocktail added to the pellet. Cells were then broken open using a ribolyser. The resulting lysate was transferred to clean tubes by using a heated needle to puncture a hole in the bottom of the tube, placing the punctured tube into another tube and briefly spinning in a centrifuge. The lysate was then spun at 16,000 × g for 5 minutes and the supernatant collected and diluted as required.

To fix the lysate to a slide, cleaned slides were treated with poly-l lysine (Sigma) for 15 minutes before being washed in ddH_2O and allowed to air dry. The lysate was then applied

to the slide, incubated for 5 minutes and the slide again washed in ddH_2O and allowed to air dry.

2.5.5 Staining with ConA

In order to differentiate between two populations of cells imaged simultaneously on the microscope one of the populations could be stained with concanavalin A (ConA), a protein that can bind to the surface of *S. pombe* cells, conjugated to Alex Fluor® 647. A 5mg/ml stock of the ConA conjugate was prepared. For use the stock was diluted to 1 mg/ml and 2 µl of the diluted ConA were added to cells suspended in 1 ml PBS which were then incubated at room temperature for 5 minutes before washing 3 × in PBS and putting on ice.

2.5.6 Treatment with DNA damaging agents

Where cells were grown asynchronously a sample was transferred to a sterile 50 ml tube and incubated with the relevant DNA damaging agent for 1-2.5 hours. For simultaneous imaging of treated and untreated cells the same volume of untreated cells were also transferred to a 50 ml tube at the same time as treated cells. Where cells were synchronised using a lactose gradient DNA damaging agents were added during the growth period from G2 to S phase for 2.5 hours. For ionising radiation (IR) treatment asynchronous cells were transferred to 50 ml tubes and exposed to 250 Gy of IR then allowed to recover at 30 °C for 30 minutes.

2.6 Single molecule localisation microscopy

2.6.1 The microscope

All single molecule localisation microscopy (SMLM) was performed on a custom built microscope (Figure 2.1). Four lasers; 405 nm (LaserBoxx Oxxius), 488 nm (Lasos, DPSS BLK 7310 T), 561 nm (Cobolt, Jive), 641 nm (Toptica TD iBeam smart), were used to give a range of fluorophore options. Each laser beam path contained a two part filter wheel fitted with 2 sets of neutral density filters (Thorlabs), one set changing in whole unit increments and the other in tenths of a unit such that they could be combined for a variety of different densities. In addition the 405 nm pathway contained a gradient filter wheel (Thorlabs) for fine adjustment of the power density. Each beam path also contained bandpass filters and ¼ wave plates to tidy up the beam followed by a beam expander. Each beam path contained a pair of mirrors to adjust the position of the beam for alignment and to direct the beam towards a shutter, each of which were remotely operatable. After the shutter the beam paths were combined using dichroic mirrors, directed into the back aperture of the microscope using a series of mirrors and a periscope and directed up into the objective (Olympus, UIS2 APON 60× OTIRF, 1.45 NA) by another dichroic mirror. Light emitted from the sample returned back through the objective and was transmitted by the dichroic. The beam path towards the electron multiplier charge coupled device (EMCCD) camera (Photometrics Evolve 512/Evolve Delta) contained a longpass filter (manually changed according to the excitation laser being used), another beam expander and an electronically powered filter wheel containing various band pass filters. The Dual View beamsplitter could be used to split the beam path for two colour imaging or could be bypassed for single colour imaging before it hit the camera. The motorised stage (Prior H117E1I4) could be remotely controlled. A white light source mounted above the objective was also available for use in finding cells on the slide and focusing the

microscope.

Imaging was controlled through μManager (Edelstein *et al.*, 2014) which allowed control of the electronic filter wheel, the shutters and the camera from a computer. Live images on the camera could be viewed at various exposures and single frame snapshots taken and saved. Multi-frame images could be acquired using the Multi-D Acquisition tool in μManager, which allowed recording of a specified number of frames at a specified time interval with a specified exposure. Alternatively scripts written in Java could be used to control all aspects of image acquisition automatically, including automatically saving sets of frames. This was particularly of use in experiments where the 405 nm activation laser was "pulsed" by opening and closing the shutter rapidly. All raw data were saved in the .tiff file format. For all experiments camera EM Gain was set to 250. Experiments were performed in an air conditioned room cooled to 18 °C.

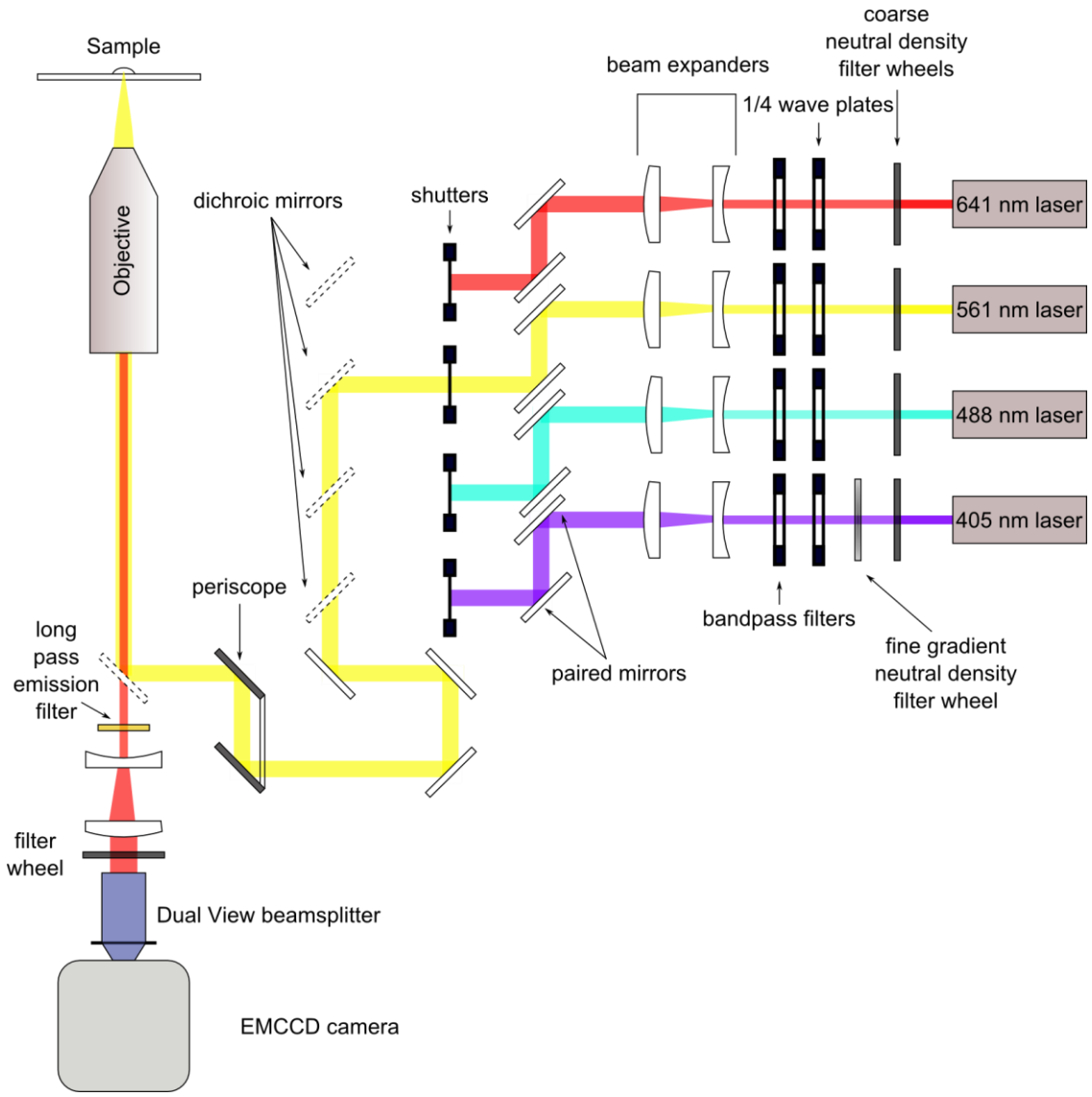


Figure 2.1 Schematic for the custom built SMLM microscope

Adapted from (Etheridge, 2015)

The microscope was set up with four different lasers, the beam paths were filtered, expanded and combined before being directed into the back aperture of the microscope. Shutters controlled which lasers illuminated the sample. Light emitted from the sample was directed into the EMCCD camera via a long pass filter, a beam expander, a filter wheel containing band pass filters and a Dual View beamsplitter.

2.6.2 The localisation software

All image processing was done using the image analysis software ImageJ/FIJI (Schindelin *et al.*, 2012) to run various single molecule localisation microscopy (SMLM) plugins created by Dr Alex Herbert (University of Sussex). The suite of plugins created by Dr Herbert covered a variety of functions from simple localisation and image reconstruction up to quantitative analysis. It also included a variety of plugins for benchmarking and simulating data in order to validate the methods being used. PeakFit is the name of the plugin which calculates the localisation position. PeakFit was used with the maximum likelihood estimation (MLE) algorithm based on Mortensen *et al.*, (2010) using a Gamma distribution for the EM-gain and a Gaussian distribution for the read noise of the camera. For MLE fitting the Gain used was 42.6 Adu/phot, the read noise was 20.64 and the Gain in ADUs was 400. Characterisation of the point spread function (PSF) of the microscope had been performed previously by Dr Thomas Etheridge (University of Sussex) and Dr Herbert.

2.6.3 Quantitative PALM

2.6.3.1 Imaging

For quantitative PALM the 561 nm laser was used at a power density of 1 kW/cm^2 and the 405 nm laser was used at varying powers adjusted manually according to activation numbers as judged by eye. Camera gain was set to 250. A bandpass filter 580/14 (Semrock) was used to exclude backscattered light and some autofluorescence.

Image acquisition was performed using either "pulsed" activation, controlled by a script written by Dr Steven Lee (University of Cambridge) and modified for our instrument, or continuous activation. For pulsed acquisitions the 405 nm laser was "pulsed" by opening and closing a shutter in the beam path for 50-100 ms depending on level of activation

required. The 405 nm pulse was then followed by an acquisition of 300 frames at 50 ms exposure, one pulse and one acquisition making up a cycle. Cycles were repeated until no more activation was seen, 405 nm power was manually adjusted between pulses when activation levels seemed to be dropping to increase the activation probability (Lee *et al.*, 2012). For continuous activation the Multi-D acquisition tool in μManager was used to acquire sets of 10,000 frames at 50 ms exposure until no more activations were seen, 405 nm laser power was increased manually using a gradient filter wheel.

2.6.3.2 Processing and analysis

For quantitative PALM the raw imaging data were fitted with localisations using the PeakFit plugin. These data were then filtered using the suggested parameters calculated for the dataset. The PC-PALM Molecules plugin was then used to trace the data (eliminate blinking and repeat localisations) and re-plot localisations in a binary image. The PC-PALM Analysis plugin was used in conjunction with the binary image to compute a pair correlation analysis for the data as described by Sengupta *et al.*, (2011), which gave a histogram containing information about long range associations. The PC-PALM Spatial Analysis plugin was used in conjunction with the binary image to perform the analysis described in Puchner *et al.*, (2013) and produced a plot of the average molecular density around a localisation as a function of distance from the localisation. The Dark Time Analysis plugin was used to calculate the time threshold needed for both clustering and spatial analysis. The PC-PALM Clustering plugin was used with localisation data to assign each localisation to a cluster based on a specified distance parameter with a variety of different algorithms. It then output a histogram of frequency of cluster size and fit this histogram with a truncated binomial to calculate values for N (the cluster size for the sample) and p (the probability of seeing a fluorophore).

2.6.4 Motion Blur

2.6.4.1 Imaging

For PALM motion blur the 561 nm laser was used at a power density of 1 kW/cm^2 and the 405 nm laser was used at varying powers adjusted manually according to activation numbers as judged by eye. Camera gain was set to 250. A bandpass filter 580/14 (Semrock FF01-580/14-25) was used to exclude backscattered light and some autofluorescence.

Before activation snapshots were taken of the cells to be imaged using white light and using the 641 nm laser with a 692/40 bandpasss filter (Semrock FF01-692/40-25). Image acquisition was performed using continuous 405 nm activation at a very low power density. The Multi-D Acquisition tool in μManager was used to acquire sets of 1000 frames at 350 ms exposure until activation was no longer seen.

2.6.4.2 Processing and analysis

All image analysis and processing was done using Fiji (Schindelin *et al.,* 2012)

The Peak Fit plugin (Herbert, 2016) was used to localise the spots from the raw data. A z-stack maximum projection was created from the raw data and used to find the positions of the nuclei for each cell, circles were drawn around each nucleus and used to create a mask (Figure 2.2). The localisation data were then filtered using this mask and the filtering parameters calculated by the plugin. The filtered data were replotted at a 10× scale as a single pixel per localisation. A scaled up version of the mask was created and, where ConA treatment has been used, each nucleus numbered or lettered according to whether it fell in a ConA+ or – cell as determined by a 641 nm laser image. The number of localisations was calculated using the Measure tool with the RawIntDen value. Values were recorded using

Microsoft Excel. Boxplots were produced using BoxPlotR ("Kick the bar chart habit," 2014; Spitzer and Wildenhain, 2016) and p values were determined using the Mann-Whitney-Wilcoxon test in R. Where p values were below the α criterion of 0.05 they were labelled with $p < 0.05$, where they were not below the α criterion they were labelled with the actual p value returned by the test. For the experiments in which the null hypothesis is that the means of the two populations do not differ and the alternative hypothesis is that the means of the two populations do differ, where the alternative hypothesis has been accepted the statistical power has been calculated using an online calculator ("Compare 2 Means 2-Sample, 2-Sided Equality | Power and Sample Size Calculators | HyLown," 2016) (see Appendix for table of results).

Figure 2.2 Method for processing motion blur data for nuclear proteins

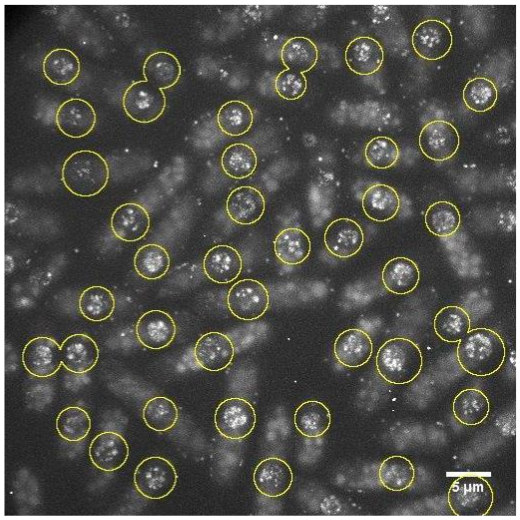
(A) A Maximum *z*-stack projection was created using Fiji and circular selections were drawn around each of the nuclei using the shift key and the circular selection tool. Nuclei that appeared out-of-focus as judged by size were not circled. These circles were then used to create a mask and the dataset was filtered using the mask and a size limit to exclude out-of-focus PSFs.

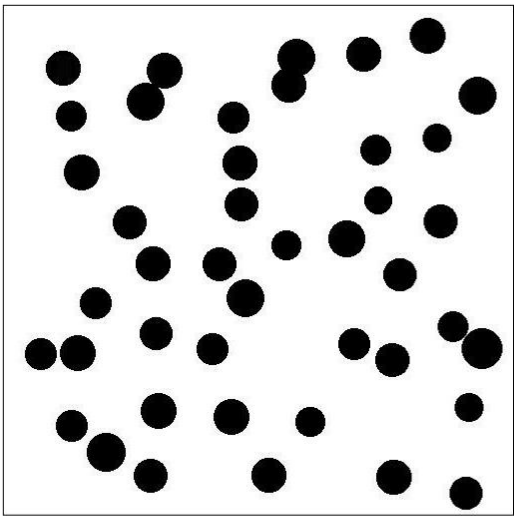
(B) The filtered localisations were then plotted as a single pixel per localisation on a 10× scale image. Where more than one localisation occurred on the same pixel the intensity of the pixel was greater (i.e., dark grey for one localisation, light grey for two and white for three). The circle tool was used to select a nucleus area and the measure command used to get the number of localisations for that area.

(C) Using the image taken with the 642 nm laser, showing the ConA conjugate dye, the nuclei could be assigned to one of two groups, treated and untreated, to allow comparison.

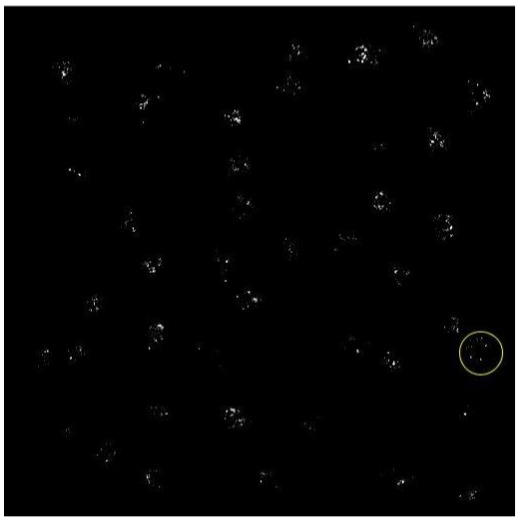
The data were not traced, so a single fluorophore that was on over more than one frame would give rise to more than one localisation, additionally the activation efficiency of mEos3.1 is thought to be ~40% (Durisic *et al.*, 2014) so this does not give an absolute count of molecule numbers but does give a relative count that can be compared between two different populations.

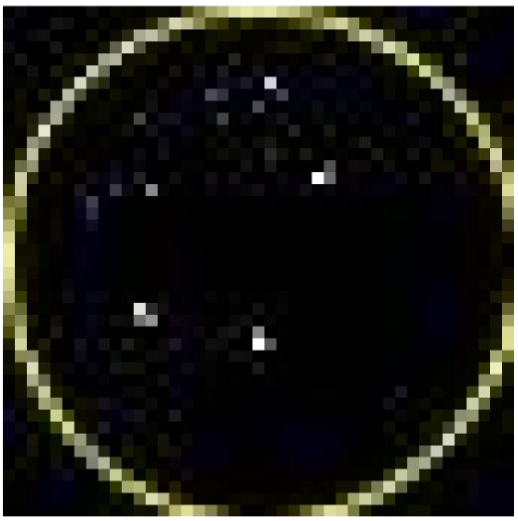
(A)



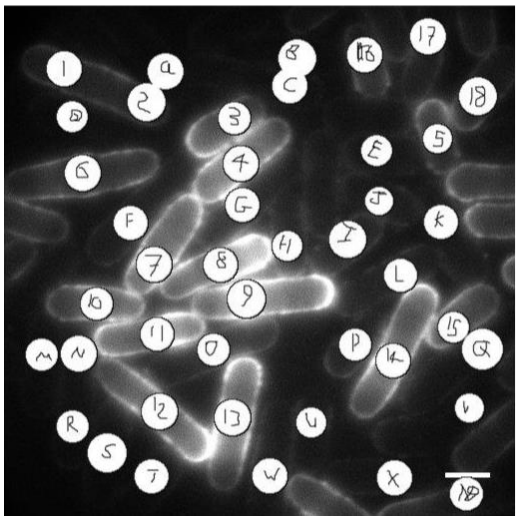


(B)





(C)



2.7 Fluorescence Correlation Spectroscopy

2.7.1 The instrument

All FCS measurements were taken on a custom built FCS instrument (Figure 2.3).

In the system used there was a 473 nm wavelength 150 mW laser (Laserglow), 561 nm 50 mW laser (CNI Laser) and a 405 nm 50 mW laser (CNI Laser). The lasers were directed into the back aperture of the microscope (Nikon TE 2000-E) via a pair of mirrors, the beam paths were combined using dichroic mirrors (Figure 2.3). Wheels containing two sets of different value neutral density filters (Thorlabs) sat in the beam paths allowing coarse adjustment of the laser power reaching the sample, in addition the beam path for the 405 nm laser contained a wheel with a continuous gradient of neutral density filter (Thorlabs) for fine tuning of the laser power. A dichroic mirror reflected laser light through the 1.4 NA objective (magnification of 100× with oil immersion) which focused the beam to a diffraction limited spot within the sample. Fluorescence emitted from the sample was transmitted by the dichroic. For mEos and other fluorophores excited by the 561 nm laser a 561 nm longpass filter (Semrock) excluded scattered laser light. A prism directed light towards either the eyepiece of the microscope stand, a complementary metal-oxide-semiconductor (CMOS) camera (USB 2.0 CMOS Camera, 1280 x 1024, Color Sensor, Thorlabs) (to check alignment and find sample cells) or to the single particle avalanche photodiode (SPAD) (Perkin Elmer). The sample could be illuminated in wide field from above by a white light lamp to enable visualisation of a cell when viewing the sample on the CMOS camera, allowing it to be positioned where the laser volume would be. A total internal reflection fluorescence (TIRF) lens on a flip mount was added at the back aperture to the microscope. This allowed the laser illumination to be switched between the highly focused, diffraction limited volume needed for FCS measurements and a de-focused mode that would allow photo-bleaching of fluorescent molecules throughout the cell.

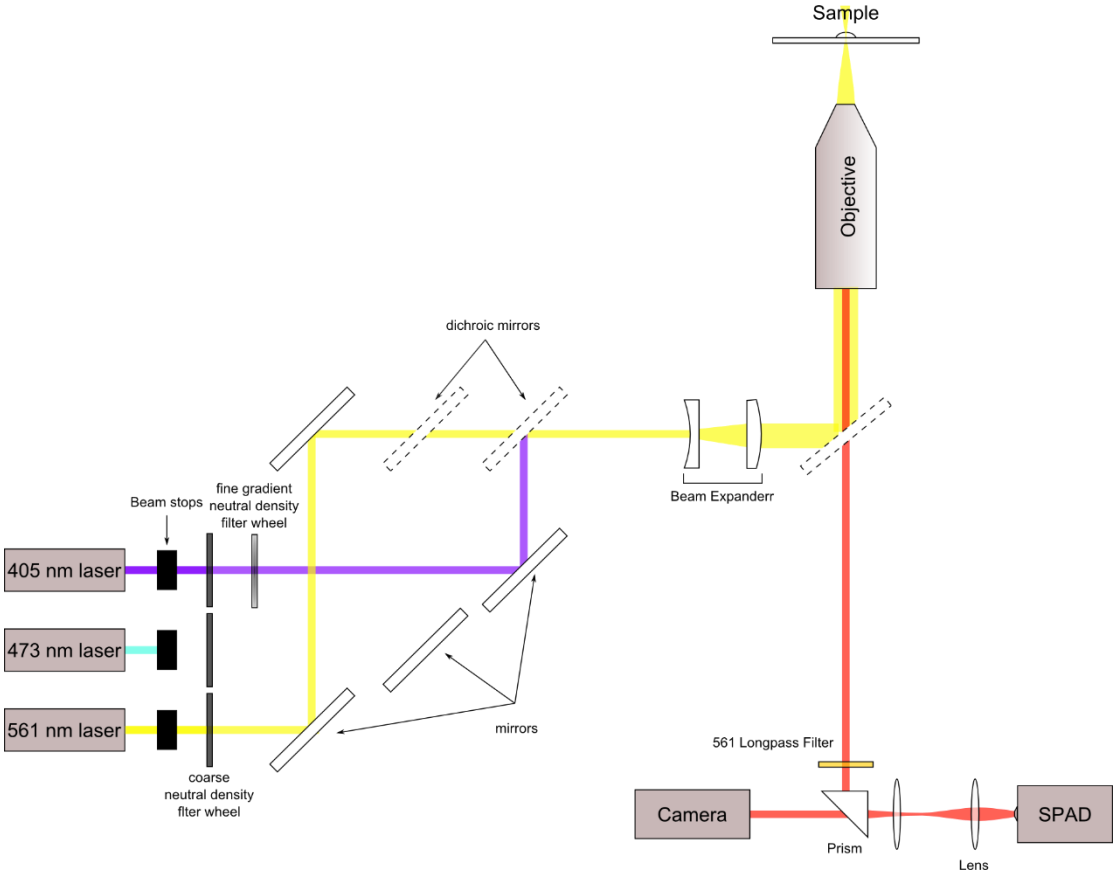


Figure 2.3 Schematic for custom FCS set up

The system used had three different lasers, a 561 nm, a 405 nm and a 473 nm. Each beam path was directed by a pair of mirrors then combined with the other beam paths using a dichroic mirror. Each laser path contained a set of neutral density filters of varying densities mounted in a wheel to control the laser power, the 405 nm laser path contained an additional gradient filter wheel for fine control. The combined laser path was directed through a beam expander and then into the back aperture of the microscope and focused into a small spot within the sample. Fluorescence from the sample passed through a 561 nm long pass filter and was directed towards either the eyepiece (not shown), a CMOS camera or the SPAD by use of a prism that could be rotated electronically. The position of the SPAD aperture means out-of-focus fluorescence is excluded from the detector.

2.7.2 Measuring a dye in solution

Organic dyes were purchased in powdered form and a stock solution created in the manufacturer recommended solvent. Rhodamine B (R6626 Sigma) was diluted to concentrations of 100 nM, 10 nM and 10 pM in water, purified mEos3.2 (prepared by Dr Adam Watson) was diluted to a concentration of 0.05 mg/ml in water. Measurements of purified dyes were taken by applying 10 µl of the diluted dye to a clean glass slide with a self-adhesive well (Biorad) attached to avoid evaporation of liquid during the experiment. Measurements were taken using a 561 nm laser at varying powers, controlled using a combination of neutral density filters (Thorlabs).

Rhodamine B labelled dextrans (ChemQuest) were diluted to stock solutions of 20 nM in water. These stock solutions were used to produce the different proportion mixtures on clean glass slides with a self-adhesive well (Biorad) attached to avoid evaporation during the experiment. Measurements were taken as for purified Rhodamine B.

2.7.3 Measuring inside cells

Live cells were prepared as described above. To ensure the laser beam would be focused inside a cell light was directed towards the CMOS camera in the absence of the 561 longpass filter and the position of the beam in the viewing window was marked. The cells were then examined on the CMOS camera using white light illumination and the stage re-positioned such that the beam mark was inside the cytoplasm and away from the cell walls.

Once the cells were positioned the excitation laser beam path was opened, light directed towards the SPAD and the intensity fluctuations could be followed on the computer. The point at which the excitation beam was in-focus was assumed to be that where the intensity peaks were highest above the background. Very low powers of 405 nm wavelength light

were used to activate continuously at low levels. The 405 nm spot was very slightly misaligned from the 561 nm spot so as not to activate molecules in the middle of the detection volume. For each cell 100× 3 second acquisitions were collected.

2.7.4 FCS measurements and autocorrelation calculation

Laser power was adjusted using a two sets of neutral density filters (Thorlabs) mounted in a filter wheel.

Signal from the SPAD was monitored using PMS-300, PMS-400 and PMS-400A, 800 MHz Gated Photon Counters/Multiscalers (Becker & Hickl GmbH, Germany). The autocorrelation function was calculated using Flex-2ksb (Correlator.com) at the same time as acquisition.

In order to extract diffusion data from the autocorrelation function it is necessary to fit it with a curve that models the expected system. The equation for the curve has to model what is expected in the data; so for a single species with a single diffusion coefficient a single component fit (*2-1*) was used. However if there are suspected to be two different populations in the sample with different diffusion coefficients it is necessary to use a two component curve to fit the data (*2-2*) (see 4.1.1.5).

$$\boldsymbol{G(t) = a + \frac{1}{N}\left(\frac{1+t}{\tau}\right)^{-1}} \qquad \textit{2-1}$$

$$\boldsymbol{G(t) = a + \frac{1}{N}\left(f\left(\frac{1+t}{\tau_1}\right)^{-1} + (1-f)\left(\frac{1+t}{\tau_2}\right)^{-1}\right)} \qquad \textit{2-2}$$

Where N = the average number of fluorescent molecules in the volume, f = the fraction of component 1, t = a time value and τ = the diffusion time. It is possible to fit up to a three

component curve but where there are more than three species, or where the diffusion coefficients of the species are very close, it can be better to use the maximum entropy distribution fit, which will give distinct peaks if there are multiple distinguishable species or show a broad distribution if there are multiple species with close together diffusion coefficients.

A series of 100× 3 s acquisitions were taken and their autocorrelation curves saved. Using a program written by Dr Mark Osborne (University of Sussex), based on a paper by Ries *et al.*, (2010), an average curve was created and the squared difference (dG) calculated for each individual curve. The curves were then ranked from minimum to maximum dG and the mean dG calculated, then a threshold of 10× the minimum dG was used to reject curves that deviated too far from average. A new averaged curve was then plotted from the curves that were kept and the data for this curve saved. The purpose of this averaging of very short acquisitions was to avoid giving undue weight to artefacts and anomalies in the data.

This average curve was then loaded into a piece of software called QuickFit (Jan Wolfgang Krieger and Jorg Langowski, 2015). This software had a conventional FCS curve fitting plugin, which would fit the curve with one, two or three components plus zero, one or two non-fluorescent components, and a maximum entropy fitting plugin. To fit with the maximum entropy plugin an estimate of the triplet (non-fluorescent state) time constant and the triplet fraction were needed, which were obtained from the conventional fit. The conventional fit plugin allowed a selection of different fitting methods, the Levenberg-Marquard fit with box constraints was used as this gave the highest R^2 values on average across a set of samples and the lowest Akaike information criterion score. The values output include N which is the average number of fluorophores in the detection volume at any one time, a diffusion time τ and fraction for each of the populations fitted, a τ and

fraction for the triplet state and, if a spot radius was entered, a calculation of the diffusion coefficient for each species. The major output of the maximum entropy plugin was a distribution plot and a plot of the curve fit along with the residuals.

3 Quantitative PALM for cytoplasmic proteins in *S. pombe*

3.1 Introduction

As discussed in the Introduction (1.2.5.3) one biological application of photo-activated localisation microscopy (PALM) is using it to quantify proteins of interest. This aspect has not been tested in the model organism *S. pombe*, nor has much work been done on quantifying cytoplasmic proteins. In this chapter the theory of quantitative PALM is introduced along with a discussion of how it has been applied in other biological questions. Subsequently the methods used to create fluorescent cytoplasmic oligomers of known stoichiometry, sample preparation and data analysis are described. The results of simulating different oligomers with varying parameters are presented along with a comparison to experimental data using artificial oligomers of mEos3.1 in cells and in a lysate, and artificial oligomers of mEos2 in cells. Then the implications of the results for the study of biological proteins are discussed.

3.1.1 Counting with single molecule localisation microscopy

Fluorescence microscopy revolutionised biological understanding of cells because its increased contrast allowed a greater resolution than previously achievable whilst visualising living cells, but it was still limited by the diffraction properties of light. Super-resolution microscopy helped to break this limit, giving resolution of tens of nanometres, down from hundreds, allowing visualisation of some of the cell's most delicate structures (see Introduction 1.2.5.3 for examples of biological applications of super-resolution microscopy). However it is not only macro molecular structures that are of interest but

also the stoichiometry of these complexes and the absolute numbers of particular proteins that are recruited to certain complexes. Before the developement of super-resolution microscopy quantitative studies of proteins *in vivo* were only achievable in bulk (e.g. by fluorescence intensity changes) or *in vitro* using biophysical techniques such as Cryo-EM. Since the conditions needed to perform Cryo-EM, or any other *in vitro* biophysical technique, differ from the conditions in a cell, the results may not match what would be observed *in vivo*. Within a cell the conditions under which a protein is expressed, folds and can perform its function can be very specific and, ideally, molecular biologists would like to observe proteins under these native conditions. Single molecule localisation microscopy (SMLM) techniques (such as PALM and stochastic optical reconstruction microscopy (STORM) (Betzig *et al.*, 2006b; Rust *et al.*, 2006b)) achieve their heightened resolution through temporal separation of each imaged molecule. In PALM photo-activatable and photo-switchable fluorophores are used that can be "activated", therefore becoming visible on the microscope, and can also be photo-bleached, a process which renders them permanently invisible. If it is assumed that each molecule is activated once, imaged and photo-bleached then it should be possible to calculate absolute numbers of fluorescent molecules in a diffraction limited spot. In addition, since these techniques can be performed using photo-activatable fluorescent proteins (PA-FPs), a one-to-one ratio of fluorophore-to-target-protein can theoretically be achieved by creating a PA-FP fusion with the protein of interest. Essentially the number of fluorescent molecules counted would be equal to the number of molecules of the protein of interest.

Attempts to measure molecular stoichiometry using conventional fluorescence microscopy include step-wise photo-bleaching (Ulbrich and Isacoff, 2007). All of the subunits of a protein of known stoichiometry were tagged with monomeric eGFP and when all of the subunits were excited simultaneously it was observed that the emission intensity reduced in a step-wise, rather than in a continuous fashion. The number of steps corresponded to the

number of subunits in the oligomer. This study was performed in a membrane using total internal reflection fluorescence (TIRF) illumination which works by using an angle of incidence such that an evanescent wave of light is created at the sample-slide interface (see Introduction 1.2.4), which only penetrates a short distance into the sample. This means that fluorophores which are out-of-focus are not excited and therefore that the background is reduced. Since the evanescent wave can only penetrate a limited distance into the cell this technique is excellent for improving the signal-to-noise ratio (SNR) when studying proteins on the cell surface or in the membrane, but is not generally applicable to cytoplasmic or nuclear proteins. Step-wise photo-bleaching was found to work optimally at densities of 20 – 200 oligomers in a 15 × 15 μm^2 area, a density at which sufficient oligomers could be sampled but where the chance of two overlapping was small. Another limitation of this technique is that it becomes difficult to distinguish the intensity steps in oligomers of 5 or more subunits, although estimates for the number of subunits can still be made based on the size of bleaching steps and the total starting fluorescence. Ulbrich and Isacoff stated that only a small number of single molecule measurements are needed to provide data on the number and mixture of subunits.

SMLM has been used to deduce the stoichiometry of ryanodine receptors in the membranes of rat cardiomyocytes. (Baddeley *et al.,* 2009). Ryanodine receptors have been shown to form crystalline arrays in lipid bilayers with a repeating subunit structure; each occupies an area of approximately 29 × 29 × 12 nm^3. They estimated the numbers of receptor proteins in an array based on the diameter of the array and the space the receptor protein was assumed to occupy. This technique was able to give quantitative results but is only useful for cases where the subunit has a structure of known dimensions and forms a repeating pattern.

SMLM has also been important for determining that T-cell antigen receptors and the linker

for activation of T-cells (LAT) cluster into islands or lipid rafts in activated T-cells. The enhanced resolution afforded by SMLM allowed Lillemeier *et al.* to apply the Ripley's K function to images of LAT and T-cell antigen receptors in both activated and inactivated T-cell membranes (Lillemeier *et al.* , 2010). The Ripley's K function allows a comparison of the density at different distances with the expected density given a uniform and random distribution. The study showed that the distribution of LAT and T-cell antigen receptors differed from that predicted for a random distribution. The distances involved would have been too small to register with diffraction limited fluorescence microscopy. The Lillemeier approach gives some information about the way in which proteins are associating but no quantitative data on how many are associating.

Another study on membrane proteins used the pairwise correlation function in combination with PALM (pairwise correlation PALM or PC-PALM) to demonstrate that associations between membrane proteins were occurring. Moreover, they were able to estimate roughly how many molecules were present in a group and over what distance the molecules associated (Sengupta *et al.* , 2011). PC-PALM compares an image to itself using a spatial auto-correlation. If an image is duplicated and the two overlaid, the correlation calculated for the images would be 1 as they are completely similar. When the images are offset by a distance the correlation decreases; the further the distance the two images are offset by the less they correlate. The two images can be offset in many different directions and distances. Because of this, the pairwise correlation function averages the correlation between the two images at each offset distance, using Fourier transforms to increase the speed at which all the correlations are computed. The pair-wise correlation function gives a quantification of the probability of finding another protein a distance *r* away from a protein compared to the expected random distribution. The total pair-wise correlation function for all peaks in a PALM image can be represented as:

$$g(r)^{peaks} = (g(r)^{centroid} + g(r)^{protein}) * g(r)^{PSF} \quad 3\text{-}1$$

Where * represents convolution. The convolution of the correlation for the centroid ($g(r)^{centroid}$) with the correlation for the point spread function (PSF) ($g(r)^{PSF}$) is defined as $g(r)^{stoch}$ and quantifies the correlation arising from multiple appearances of the same molecule, and the convolution of $g(r)^{PSF}$ with the protein correlation function for $r>0$ ($g(r)^{protein}$) represents the spatial distribution of the proteins. For a randomly distributed protein $g(r)^{protein}$ is ~1 meaning the total pair-wise correlation function for a randomly distributed protein is.

$$g(r)^{peaks} = g(r)^{stoch} + 1 \quad 3\text{-}2$$

If the distribution is non-random, then $g(r)^{protein}$ will be greater than 1. By plotting the autocorrelation function $g(r)$ against the distance r a curve is produced that can be approximated by an exponential decay. If there is only a single object the pairwise correlation function produces a simple curve of decreasing correlation. If there are two objects the same curve is produced but with twice the height. The farther apart the two points are, the broader the spread of the curve (Figure 3.1). The PC-PALM curves are normalised to the total object count.

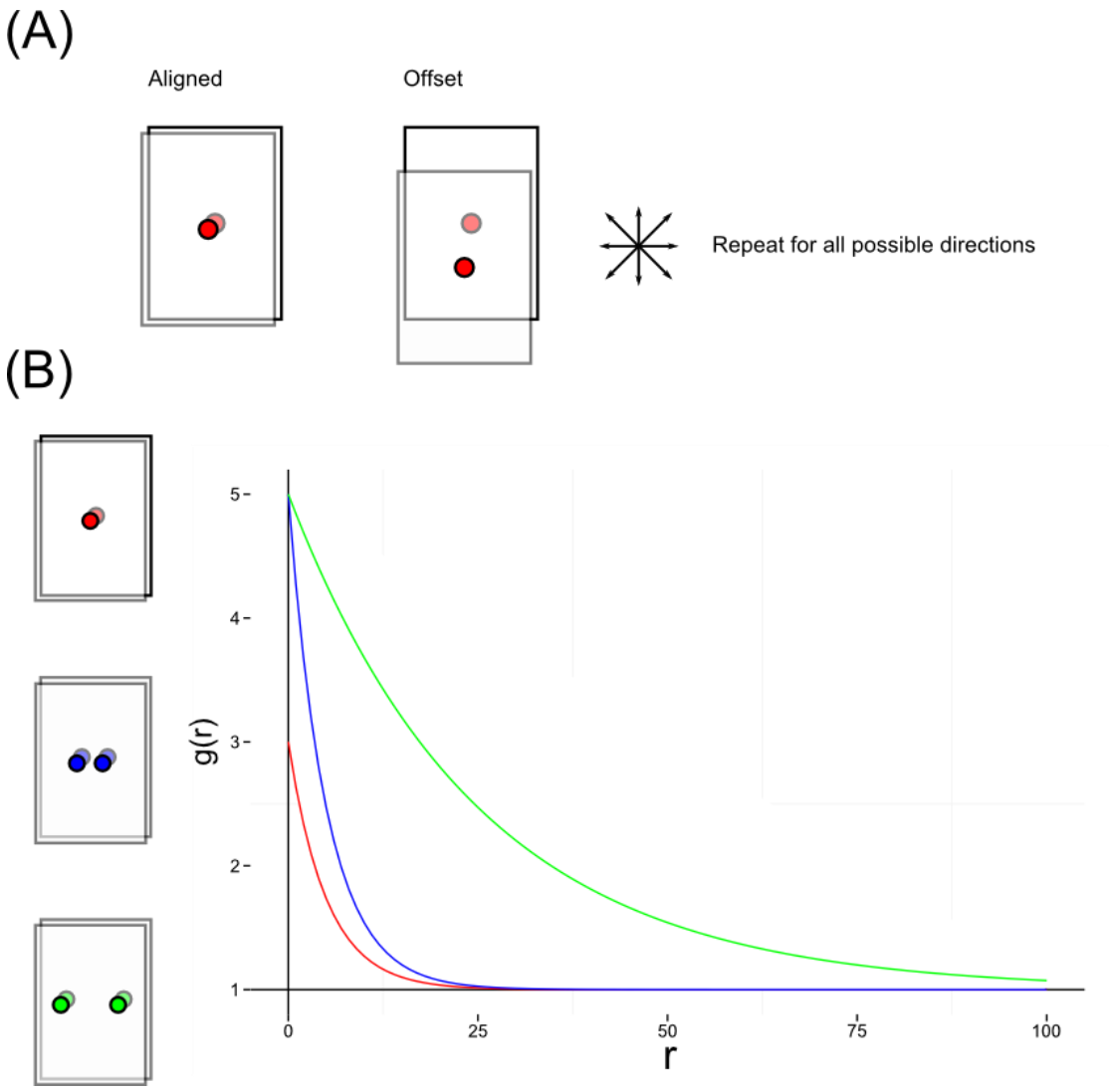


Figure 3.1 Graphical representation of the pairwise correlation function

(A) Two duplicated images aligned so that the correlation equals 1 and offset so that the correlation does not equal 1.

(B) Examples of curves that might be seen from a pairwise correlation experiment. The red image contains a single object and gives a simple exponential decay from its autocorrelation. The blue image contains two objects; the peak for its correlation curve is twice as high as the red. The green image also contains two objects but they are spaced farther apart than in the blue image. This gives a curve that peaks at the same height as the blue, but which is more broadly spread.

From the height and spread of the curve it is possible to estimate the number of molecules and their distance from each other by fitting a model to the curve. There are two well understood models for the pairwise correlation curve: The fluctuation model (equation *3-3*), where molecules are only roughly associated, and the emulsion model (equation *3-4*), where every localisation exists within one of a series of circles with no localisations between circles (Figure 3.2).

$$g(r) = 1 + A \times r^{-1/4} \times \exp\left(-\frac{r}{\xi}\right) \qquad 3\text{-}3$$

$$g(r) = 1 + A \times \exp\left(-\frac{r}{\alpha}\right) \times \cos\left(\frac{\pi r}{2r_o}\right) \qquad 3\text{-}4$$

Where A is amplitude of the correlation extrapolated for $r=0$, a is a measure of the coherence length between circles, r_0 is the average circle radius and ξ is the correlation length, a measure of the domain size of the protein cluster (Veatch *et al.*, 2012). The models work well for proteins that associate at distances that are much greater than the resolution of the localisation technique. At short distances, however, the pairwise correlation curve is corrupted and this can be modelled by the convolution of the exponential component of the $g(r)^{peaks}$ equation with the $g(r)^{PSF}$, i.e. the correlation is blurred by the uncertainty of the localisations. The error at short distances may also be due to pixelation noise in the super-resolution image that is used for calculating the *g(r)* curve because each localisation is rounded to the nearest pixel. This causes noise where the distance between localisations is short because localisations may not be recorded at the correct distance from each other. At larger distances, the rounding effect is a smaller fraction of the distance between any two localisations. In addition, the number of samples at low offset values is small resulting in a large standard error for the *g(r)* value at low offset distances. For PALM the resolution is typically 10 – 30 nm and the fluctuation model has been demonstrated to show associations in the range of 100 – 200 nm (Sengupta *et al.*,

2011). This means PC-PALM is not appropriate for tightly associated proteins, clustering at tens of nanometres, distributed randomly throughout the cytoplasm or nucleus of a cell. Additionally, the PC-PALM approach was designed with membrane proteins in mind and so only works for two dimensional data. One major reason that membrane proteins have been so well-studied is because they can be imaged using TIRF, which helps reduce background fluorescence from the rest of the cell, but which cannot be used to study proteins in the centre of a cell. Thus, it is unclear how generally applicable these methods are to studying non-membrane proteins.

Ricci *et al.*, (2015) used STORM to quantify the clustering of nucleosomes *in vivo* in the nucleus of various mammalian cell lines. By comparing the number of localisations they saw to an *in vitro* STORM experiment, in which they controlled the size of nucleosome clusters, they were able to estimate the number of nucleosomes in each cluster. They were able to use this technique to demonstrate that smaller nucleosome clusters correlated with regions of active chromatin, that large clusters correlated with areas of heterochromatin and that there was a correlation between clutch size and cell pluripotency. The nature of STORM means that the same molecules are localised repeatedly and they found that the labelling ratio of their antibody was greater than 1:1, meaning that this technique cannot give absolute numbers of nucleosomes, but it is useful for comparative quantification. It also relies on being able to compare localisation numbers to an *in vitro* standard, something which is only achievable with complexes which are stable outside of the cell.

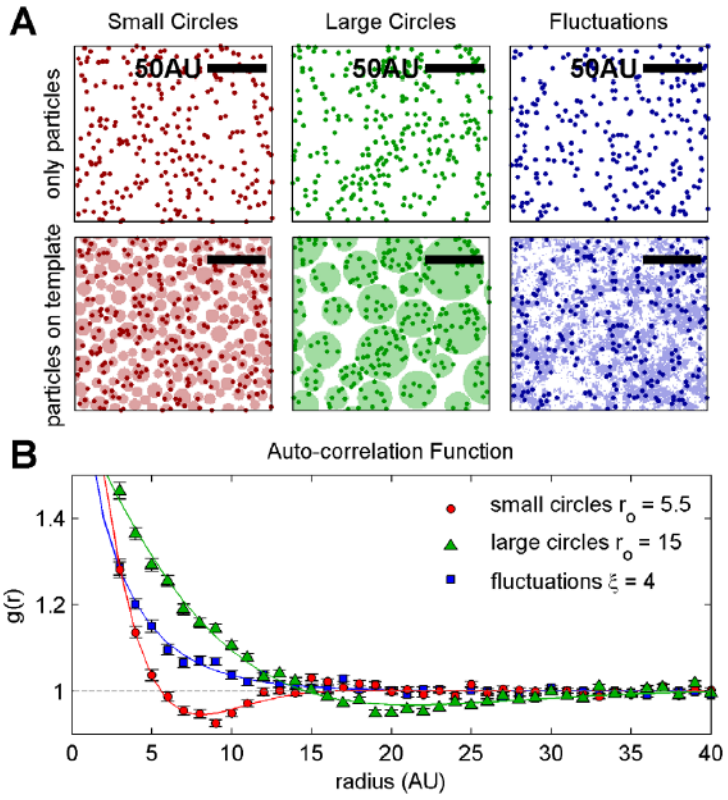


Figure 3.2 The models used with PC-PALM

Taken from Veatch *et al.*, (2012)

Two ways of applying the emulsion model, with small circles and with large circles, and the fluctuation model. The expected *g(r)* for a completely random distribution is 1.

(A) Particles distribution were simulated according to either the emulsion or the fluctuation model, the emulsion model was plotted with small circles and with large circles. The radii of the circles are plotted in arbitrary units (AU) and each particle has a radius of 2 AU. The small circles have radii between 4 AU and 8 AU, large circles have radii between 10 AU and 30 AU. The fluctuation circles were produced by simulating an Ising model at $T = 1.075\ Tc$, where *Tc* is the critical temperature and the predicted correlation length (ξ) is 4 AU.

(B) The autocorrelation function for each of the three distributions. For the emulsion models the curve is fit to equation *3-4* and both emulsion models dip below $g(r) = 1$. Whereas the fluctuation model does not dip below $g(r) = 1$ and fits to equation *3-3*.

Another obstacle to using PALM quantitatively that has become apparent is that the "blinking" properties of many of the organic dyes and PA-FPs used for PALM lead to over-counting (Annibale *et al.* , 2011). Annibale *et al.*, (2011) described a method of single molecule counting that took the blinking properties of PA-FPs into account. They did this by combining localisations from each blink together based on a time threshold. They calculated this threshold for mEos2, which is one of the PA-FPs frequently used for PALM, and described how this could be applied to other PA-FPs. Lando *et al.* , (2012) went on to use this method to look at a protein that forms a single focus in the nucleus of cells (Cnp1). They were able to count the number of Cnp1 molecules and, because it is found in a single focus, they were able to discard any localisations outside of the focus. They used this technique to follow the levels of Cnp1 throughout the cell cycle and were able to establish that it is only deposited during the G2 phase.

Gunzenhäuser *et al.*, (2012) combined Annibale *et al.*'s thresholding method with TIRF imaging to look at the HIV-Gag protein forming clusters at the plasma membrane. Using SMLM allowed them to study clusters of a range of sizes at nanoscale resolution, which had not previously been possible. Standard fluorescence microscopy lacked the resolution to determine cluster stoichiometry and the sensitivity to detect the small clusters. They used this SMLM approach to study how different fluorescent labels affected the assembly of the HIV-Gag protein clusters.

Lee *et al.*, (2012) also looked into how the blinking properties of PA-FPs can lead to over-counting in quantitative PALM. They compared two of the commonly used PA-FPs, Dendra2 and mEos2, and proposed a kinetic model to describe blinking, a photo-activation regime to reduce simultaneous activations (and thus undercounting) and a method for counting using PALM. The blinking properties of Dendra2 and mEos2 were initially assessed *in vitro* using biotinylated mEos2 and Dendra2 immobilised on streptavidin-coated

glass. It was found that Dendra2, although dimmer than mEos2, blinked fewer times on average per molecule than mEos2 and photo-bleached faster. As such, it was thought to be a better fluorophore for quantitative PALM. The kinetic model assumed four states for the PA-FPs; non-active, active, dark and photo-bleached and described the probability of the number of blinks, the fluorescent on time, the time spent in the dark state and the time until photo-bleaching. The proposed activation regime used continuous modulation of the 405 nm laser starting from a very low power such that the activation events followed a Fermi function, a smooth approximation of a uniform distribution.

To tackle the problem of quantitative SMLM it is advantageous to combine analysis of the blinking properties such as that performed by Annibale *et al*, and spatial distribution analysis such as the Ripley's K function used by Lillmeier *et al*, or the pairwise correlation function used by Sengupta *et al*. In 2013 a paper was published that used this combined approach: Puchner *et al*., looked at spatial distribution, correcting for blinking properties, to determine absolute numbers of membrane proteins using artificial oligomers consisting of one, two or three linked units of mEos2 as a calibration (Puchner *et al*., 2013) (Figure 3.3 A). The observed density at different distances was computed using a function similar to the Ripley's K function used by Lillemeier *et al*, that gave a histogram which should be flat for randomly distributed molecules but show an elevation at low distances when there were multiple molecules in the same place, either due to blinking or clustering. They used parameters of 150 nm and 2.66 s to trace together blinking events. They were able to show that for the monomeric construct, after correcting for blinking, the density histogram showed a plateau (Figure 3.3 C), whereas for dimer and trimer it showed a peak at short distances, indicating that the molecules were clustering. They stated that the width of the density peak was the precision of localisation and they used the value to set the distance for a clustering algorithm. The clustering algorithm was used to assign localisations to clusters based on a specified distance parameter. A histogram of the resulting frequencies of

cluster size was plotted and fitted with a binomial curve. The *N* of the binomial equation gave the number of fluorophores in the cluster and they were able to count artificially produced monomers, dimers and trimers of mEos2 expressed in a membrane. (Figure 3.3 B).

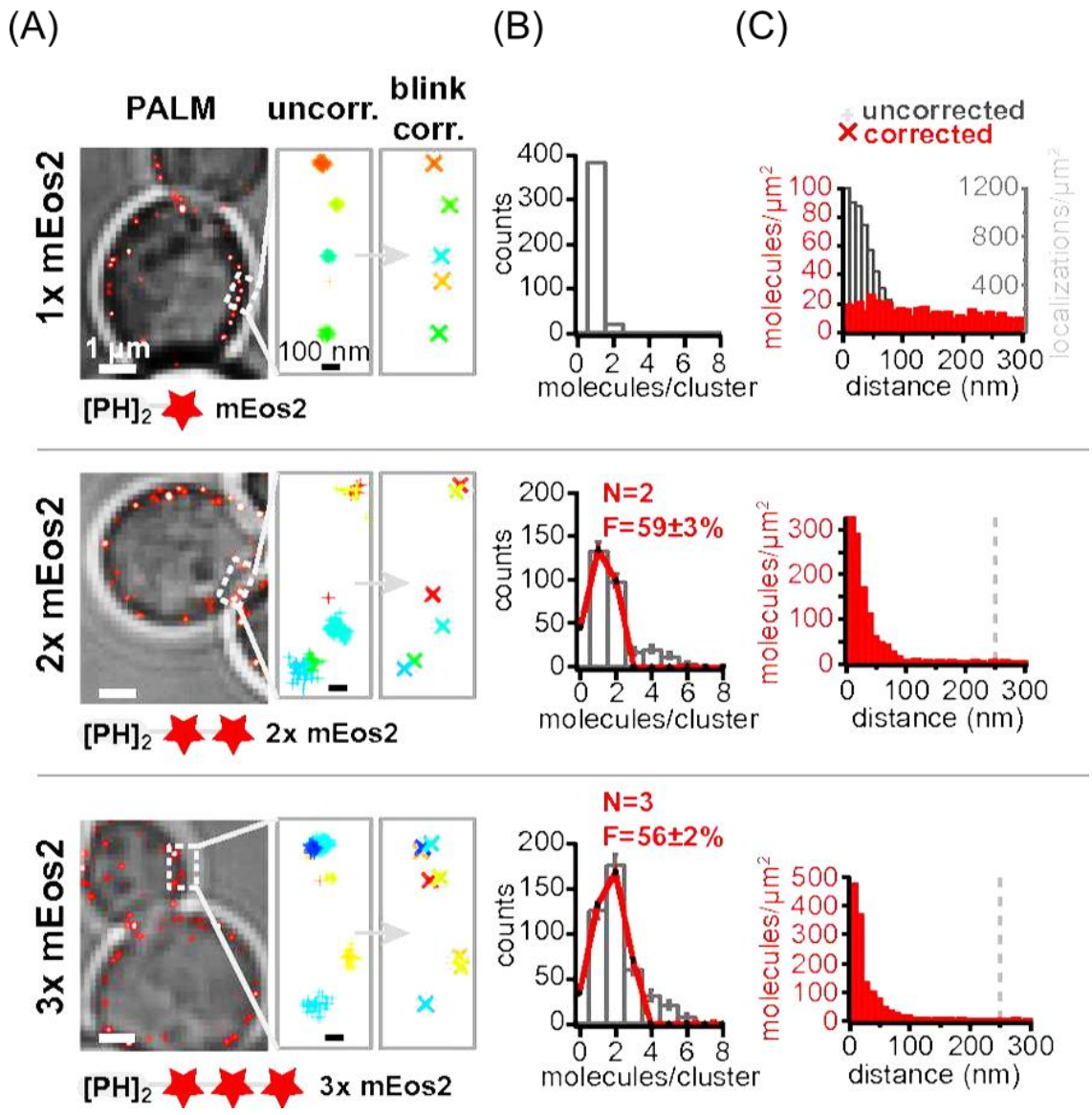
Figure 3.3 Intracellular calibration of super-resolution microscopy allows counting of biomolecules in sub-diffraction limited structures

Taken from Puchner *et al.*, (2013)

(A) Puchner *et al.*, (2013) created constructs that consisted of one, two and three repeats of mEos2 with a membrane localisation signal. They performed a PALM experiment and corrected their data for blinking by tracing together localisations within 150 nm and 2.66 seconds of each other.

(B) They then ran a clustering algorithm, produced a histogram of the sizes of the clusters they found and fitted it with a binomial distribution with an N equal to the number of mEos2 repeats and found an F equal to the reported photo-activation efficiency for mEos2.

(C) They also computed the observed density/expected density for their data both before and after correcting for blinking. For the monomer the density correlation output forms a plateau because it is completely randomly distributed. For the dimer and trimer the density function gives a peak at short distances because it is not randomly distributed (i.e. the fluorophores cluster together because they are multimeric). The width of the peak was said to be the precision of their localisation and was used to set the clustering threshold distance.


(A)
(B)
(C)
PALM
uncorr.
blink corr.
1x mEos2
1 µm
100 nm
[PH]2 mEos2
counts
molecules/cluster
uncorrected
corrected
molecules/µm²
localizations/µm²
distance (nm)
2x mEos2
N=2
F=59±3%
[PH]2 2x mEos2
3x mEos2
N=3
F=56±2%
[PH]2 3x mEos2

3.1.2 Objectives: Assessing quantitative PALM in *S. pombe*

The objective of this work is to assess the potential advatages and limitations of quantitative PALM for studying cytosolic or nuclear (as opposed to membrane) proteins in the model organism *S. pombe*. In particular it is intended to establish whether this technique could be used to asess the stoichiometry of $Cdc22^{R1}$, the cytoplasmic subunit of the ribonucleotide reductase complex. To this end, it was necessary to find a protein that formed complexes or clusters of a known stoichiometry, at a diffraction limited distance, *in vivo*. No endogenous proteins were found for which the stoichiometry was established and which were easily tagged, so the other option was to introduce an engineered protein containing a defined number of repeats of a PA-FP. This was done by creating constructs containing one, two or three copies of the PA-FP mEos3.1. These were imaged using PALM and then attempts were made to count one two or three molecules in a single diffraction limited spot. At the same time computer simulations were used to assess the conditions under which the algorithm used to count cluster sizes (based on (Puchner *et al.*, 2013)) would return the expected answer. Simulations looked at how precision, photo-activation efficiency, cluster size and cluster density affected the algorithm. Data from the oligomeric mEos3.1 constructs was then compared to the simulation data.

3.1.3 Producing constructs of a known size

Three constructs were designed consisting of a single, double or triple repeat of mEos3.1 separated by a 4×TGS linker and under the control of the uracil inducible promoter *urg1* (Watson *et al.*, 2011) (Figure 3.4 A). Dr Adam Watson (University of Sussex) designed, and produced, these constructs using the sequence for mEos3.1 with the codons optimised for expression in *S. pombe* (Forsburg, 1994) (see Appendix for sequence). The *urg1* promoter, when induced, gives a very high expression level. In addition it is a "leaky" promoter

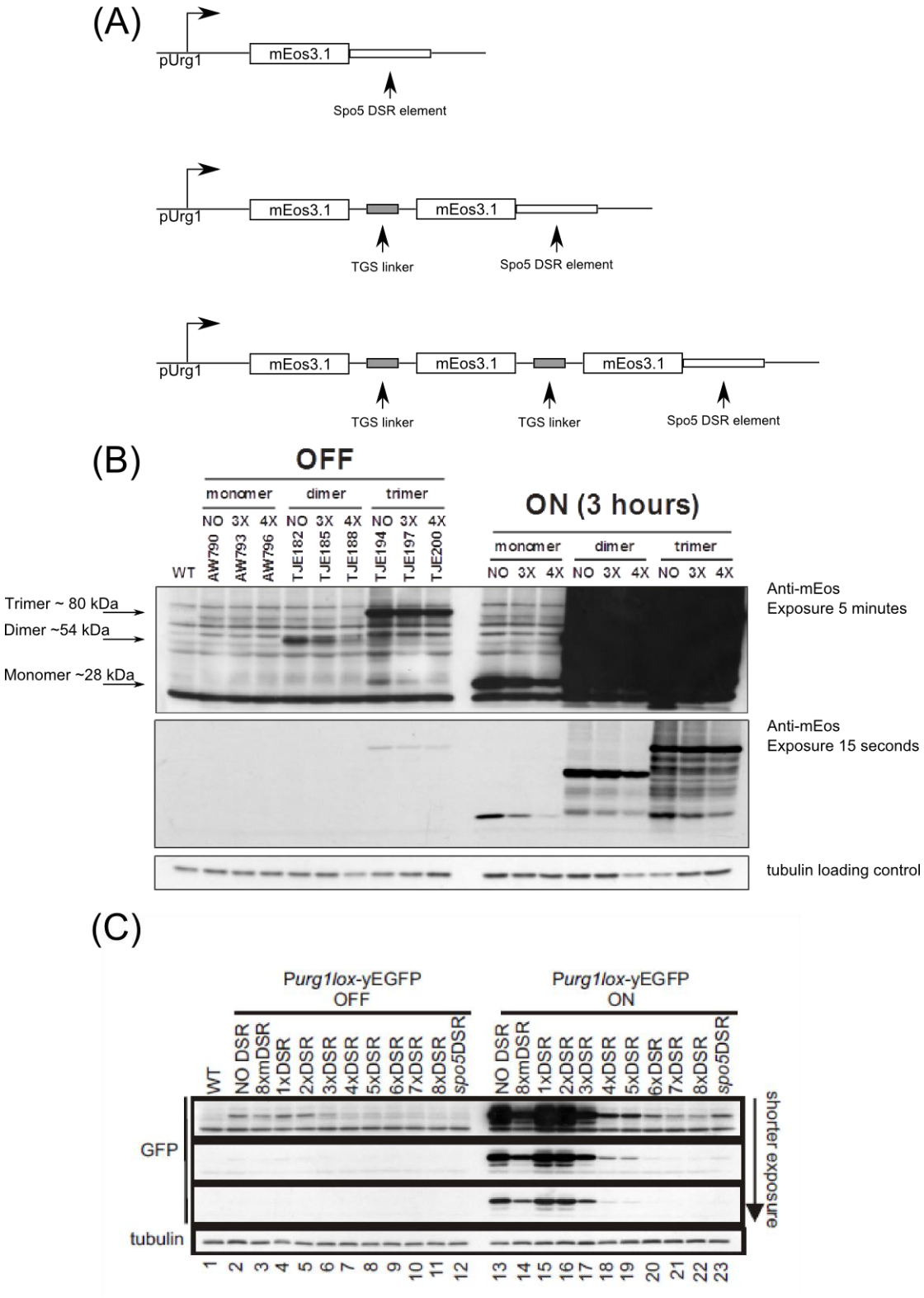
meaning that when it is not induced some expression is still seen. To regulate expression levels a sequence element called the determinant of selective removal (DSR), which reduces expression level through transcript removal (Harigaya *et al.*, 2006), was added at the C-terminal end of the gene construct. The DSR element can be repeated up to 8× and the number of repeats is proportional to the level of transcript removal, i.e. 8×DSR has the lowest expression level. In combination with the *urg1* inducible promoter these elements give a fine control over expression. The expression levels and molecular weight of these expressed proteins for these constructs were tested using a western blot of a crude cell extract (carried out by Dr Watson). Bands were observed at the correct sizes to correspond to the monomer, dimer and trimer (Figure 3.4 B). Expression levels of both induced and un-induced samples from the varying DSR levels were also tested to demonstrate the effect of increasing DSR level (Figure 3.4 C). It can be seen that as DSR level is increased expression level decreases in both the ON and the OFF samples with 8×DSR having the lowest expression. It can also be seen that the original *spo5* DSR element, an endogenous transcript removal element, which contains multiple repeats of the defined DSR sequence, gives an effect roughly the same as 6×DSR.

Figure 3.4 The mEos3.1 constructs for counting

(A) Constructs were designed by Dr Adam Watson consisting of one, two or three repeats of the mEos3.1 sequence separated by a 12 amino acid 4×TGS repeat linker sequence. Each construct was under the control of the uracil inducible promoter *urg1* (Watson *et al.*, 2011) and was followed by varying levels of DSR; a short sequence that tags transcripts for degradation (the endogenous *spo5*-DSR is indicated in the figure). The general trends is that more repeats of the DSR unit that are present the more transcripts are removed and the lower the expression levels of the protein.

(B) Western blot produced by Dr Watson (unpublished). The blot was probed with a primary antibody to mEos and shows protein extracts from both induced (ON) and un-induced (OFF) cells containing the monomer, dimer and trimer constructs with either No, 3× or 4× DSR repeats. It can be seen that the band for the dimer runs at twice the size of the band for the monomer and the band for the trimer runs at three times the size. Some level of expression can also be seen in the OFF samples, demonstrating that the *urg1* promoter is "leaky".

(C) Western blot produced by Dr Watson (unpublished). GFP has been expressed under the control of the *urg1* promoter with varying numbers of DSR (0-8 repeat copies plus the endogenous *spo5*). The *spo5* DSR repeats has the same effect as 6×DSR. In combination with the *urg1* inducible promoter these elements give a very fine control over expression. The 8x mDSR, which contains a mutated DSR sequence, shows limited activity compared to the wt DSR repeats.


(A)
pUrg1
mEos3.1
Spo5 DSR element
pUrg1
mEos3.1
mEos3.1
TGS linker
Spo5 DSR element
pUrg1
mEos3.1
mEos3.1
mEos3.1
TGS linker
TGS linker
Spo5 DSR element
(B)
OFF
monomer
dimer
trimer
NO 3X 4X NO 3X 4X NO 3X 4X
WT AW790 AW793 AW796 TJE182 TJE185 TJE188 TJE194 TJE197 TJE200
ON (3 hours)
monomer
dimer
trimer
NO 3X 4X NO 3X 4X NO 3X 4X
Trimer ~ 80 kDa
Dimer ~54 kDa
Monomer ~28 kDa
Anti-mEos
Exposure 5 minutes
Anti-mEos
Exposure 15 seconds
tubulin loading control
(C)
Purg1lox-yEGFP
OFF
Purg1lox-yEGFP
ON
WT
NO DSR
8xmDSR
1xDSR
2xDSR
3xDSR
4xDSR
5xDSR
6xDSR
7xDSR
8xDSR
spo5DSR
NO DSR
8xmDSR
1xDSR
2xDSR
3xDSR
4xDSR
5xDSR
6xDSR
7xDSR
8xDSR
spo5DSR
GFP
tubulin
shorter exposure
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23

3.1.4 Analysing the data

Following the establishment of the mEos3.1 monomer, dimer and trimer constructs an article was published using a similar approach to count single molecules using PALM. Puchner *et al.*, used monomeric, dimeric and trimeric constructs of mEos2 with a membrane localisation signal, expressed in *Saccharomyces cerevesiae* (Puchner *et al.*, 2013). Using clustering analysis and local density analysis they were able to tell the difference between their monomers, dimers and trimers *in vivo*. Dr Alex Herbert (University of Sussex), who was involved in this project as a Bioinformatician, giving technical advice and writing analysis software, wrote a series of plugins to be used with the microscopy analysis software FIJI (Schindelin *et al.*, 2012) to replicate this analysis.

The Puchner *et al.* approach involved initially grouping localisations that are close together in both space (150 nm) and time (2.66 seconds). Each group was declared to originate from the same molecule (sometimes referred to as tracing) in order to account for the blinking effect. Using a hierarchical, centroid linkage clustering algorithm the traced localisations were then further grouped into clusters using a more stringent distance threshold (50 nm) but no temporal threshold. There are a variety of ways in which these clusters can be assigned within Dr Herbert's plugin (Figure 3.5): All clustering methods share the process of iteratively joining the two closest objects. This is repeated until no more objects can be joined using a specific distance threshold. The various methods differ in how the objects are defined. In the particle single linkage method, single localisations can be added based on the distance to another single localisation. Joining localisations creates a cluster. However, established clusters are never joined into larger clusters; in essence clusters are formed by allowing single particles to join (Figure 3.5 B). The particle single linkage algorithm is useful when analysing simulated data because each join is between two single localisations and a record can be made of the distance and whether the join was made

between localisations simulated to come from the same cluster or from separate clusters (see 3.2.1). Analysis of the intra- or inter-molecular distance will be discussed later in the chapter.

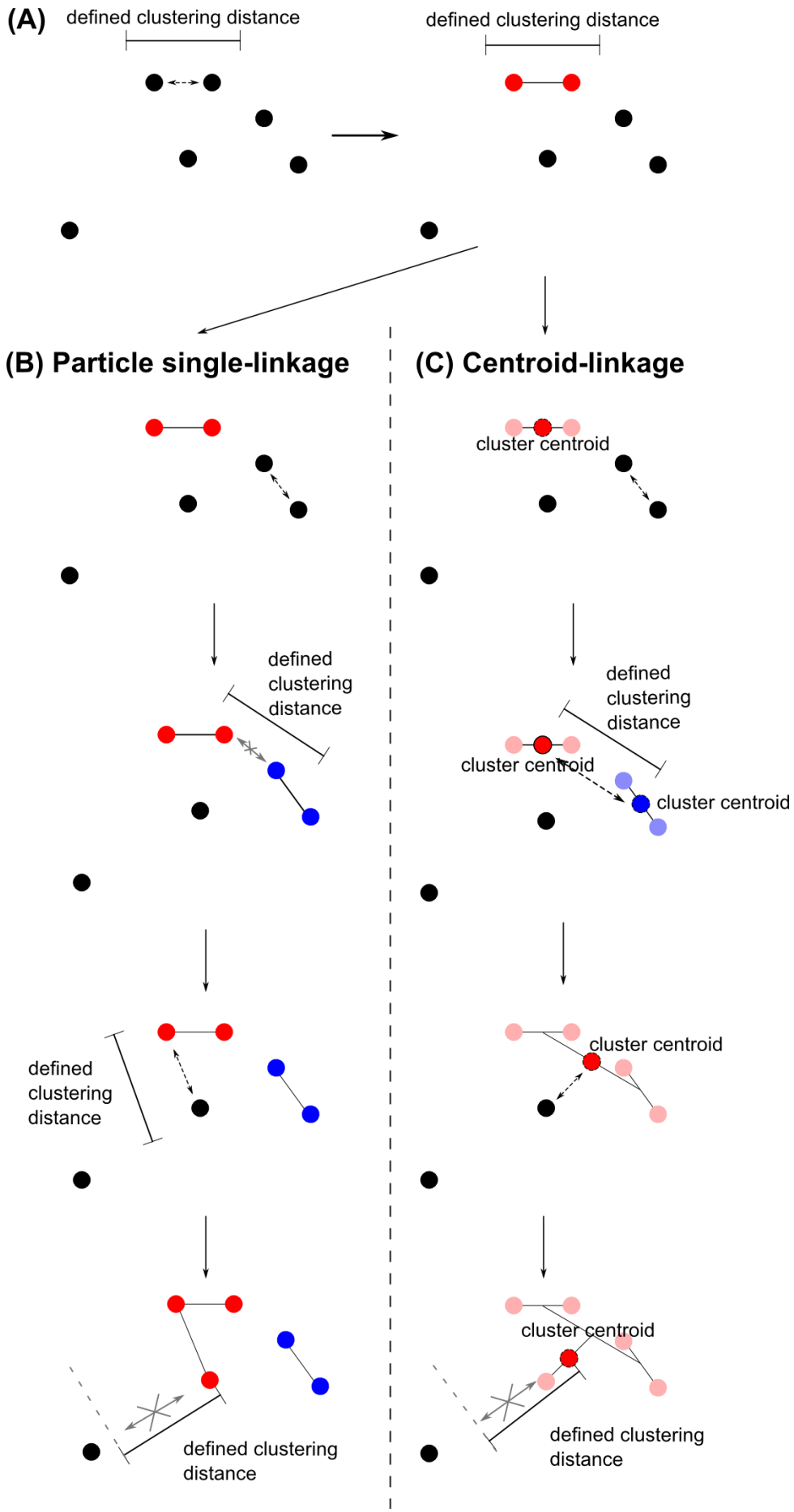
Alternatively, the same process can be followed but joins can be allowed between established clusters. This centroid linkage method re-calculates the centre-of-mass for a cluster each time a localisation is added and localisations are added based on distance to this centre-of-mass rather than the distance to the nearest localisations, which prevents the assigning of long, string-like clusters (Figure 3.5 C). This method is classic hierarchical centroid linkage clustering. However analysis of inter- and intra-molecular distance is not possible since joins may be between clusters containing many different molecules which cannot be categorised as intra- or inter-molecular joins. These different methods can potentially give quite different results in terms of cluster sizes calculated (Figure 3.5 B-C). Following clustering a histogram plot of the frequency of each cluster size can then be fitted with a binomial distribution with parameters N and p. Parameter N should be the cluster size expected (i.e. N=3 for a trimer), and parameter p is the probability of seeing a given fluorophore and should match the photo-activation efficiency.

Figure 3.5 Different methods for clustering

(A) The first step for all clustering algorithms. Two localisations, within a defined threshold distance of each other, are joined to form a cluster.

(B) The particle single-linkage algorithm for clustering. In this algorithm localisations are joined iteratively based on the threshold distance. This algorithm does not re-calculate the cluster centroid so the threshold distance is simply between two localisations, this can lead to long string-like clusters. In the "particle" algorithms a single localisation may be joined to a cluster but two clusters may not be joined together. In the final panel the un-clustered localisation is outside of the threshold distance from its nearest neighbouring localisation so is not joined to a cluster. Running the particle single-linkage algorithm on the localisations shown results in two clusters, one of $N = 2$ and one of $N = 3$, and a single un-clustered localisation.

(C) The centroid-linkage algorithm for clustering. In this algorithm a new centroid is calculated each time a localisation is added to a cluster. New localisations are added to the cluster on the basis of their distance to the centroid. In addition, this algorithm allows two clusters to be added together if their centroids are within the threshold distance of each other, forming a single cluster. In the final panel the un-clustered localisation is outside of the threshold distance from the centroid of the cluster and has no other neighbouring localisations within the threshold so it is not joined to a cluster. Running the centroid-linkage algorithm on the localisations shown results in a cluster of $N = 5$ and a single un-clustered localisation.


(A)
defined clustering distance
defined clustering distance
(B) Particle single-linkage
(C) Centroid-linkage
cluster centroid
defined clustering distance
defined clustering distance
cluster centroid
cluster centroid
defined clustering distance
cluster centroid
cluster centroid
defined clustering distance

The clustering methods use the localisations derived from the fluorophore image as input. However there are a number of potential sources of error in the data that could affect the clustering result; the blinking properties of the fluorophore, the depth of focus of the microscope, the activation power of the laser, the noise in the image, the precision of the localisation method, and the photo-activation efficiency of the fluorophore being used. Each error source is discussed below.

Blinking can, to some extent, be ruled out by using very low power "pulses" of the activation wavelength followed by a long period of image collection with the excitation wavelength such that all the "blinks" occur in a single set of frames. In this case it can be assumed that all localisations within a defined area for each set of frames come from the same molecule.

The depth of focus is a property that relates to the numerical aperture (NA) of the objective, higher NAs allow for better capture of photons, and so give a better resolution, but they can only focus to a limited depth. In the case of the 1.49 NA objective this depth is around 500 nm, about $1/6^{th}$ of the depth of an *S. pombe* cell. This means that some activated fluorophores will be out-of-focus and give a more diffuse PSF that will be harder to localise. These localisations can be excluded by filtering on the basis of PSF width (Palayret *et al.*, 2015) but there is a small risk of excluding half of an oligomer if it is on the edge of the focal plane.

If the activation power of the laser is high, molecules that are close to each other may be activated at the same time. To avoid this, the activation power must be kept very low such that only single fluorophores are activated within a certain area and the PSFs do not overlap.

The noise in the image (background fluorescence causing additional localisations) can be subtracted by looking at the noise in cells not expressing mEos3.1 or other fluorescent

proteins as described in Puchner *et al.*, (2013). This involves running a clustering algorithm on the mEos-free cells, generating a histogram of cluster sizes which is normalised by the area of cells used and subtracting that histogram (scaled to the area of experimental data) from experimental results. This mainly subtracts the over-occurrence of clusters of 1. However, this is a global correction and it cannot discriminate noise localisations from "real" localisations, instead it subtracts a rough estimate of how many localisations are likely to be due to noise. Subtracting background in this way can be inaccurate as the internal background fluorescence of the cells has been found to vary from one preparation to another. Additionally, because subtraction happens after tracing and clustering, localisations coming from the noise are included in the tracing and clustering algorithms before they can be subtracted. If the background localisations are sparse this approach is OK as the background localisations will not tend to overlap with "real" localisations. However, if the background is dense background localisations may be added to clusters by the clustering algorithm inflating the calculated cluster size.

There is inherent uncertainty in localising the position of a fluorophore from an image on a camera, this uncertainty in localisation is also called the precision. The precision is widely assumed to be normally distributed (Abraham *et al.*, 2009; Thompson *et al.*, 2002) This can be demonstrated for the system used by simulating and fitting a single localisation in the same position multiple times. The resulting localisation positions can be plotted as a histogram giving a distribution that appears approximately normal (Figure 3.6). To compare an experimental distribution with a theoretical distribution a Q-Q plot can be used (the Qs stands for quantile, the values which cut the probability distribution into intervals with equal probability) in which the quantiles of the experimental distribution are plotted against the quantiles of the theoretical distribution; when the distributions are similar the Q-Q plot will produce a straight line. By plotting the quantiles of the simulated localisation data against the theoretical quantiles of a normal distribution it can be seen that the

simulated data fits well to a normal distribution (Figure 3.6 B). The precision quoted for a system is usually one standard deviation of the distribution, where for a normal distribution there is a 68.3% probability a given localisation will fall within that distance of the absolute position (Figure 3.6 A).

The effect of the number of photons, the pixel size and PSF size on the localisation precision has been modelled by Thompson *et al*, (2002) and later Mortensen *et al*, (2010) (see Introduction 1.2.5.2). Equation *3-5* gives the precision of maximum likelihood fitting of a Gaussian PSF model where N is the number of photons in the image, b^2 is the average background per pixel, σ is the expected width of the Gaussian PSF and a is the pixel size of the detector (Mortensen *et al.*, 2010).

$$Var(\mu_x) = \frac{\sigma_a^2}{N}\left(1 + \int_0^1 \frac{\ln t}{1 + \frac{Na^2}{2\pi\sigma_a^2 b^2}} dt\right) \qquad 3\text{-}5$$

Precision varies from localisation to localisation but the average precision for the system used in this study has been calculated to be approximately 20 nm. Meaning that for a given absolute position there is a circle of radius 20 nm within which there is a 68.3% probability of a localisation falling. Therefore, a trimer of mEos3.1, assuming each unit of mEos3.1 has the same absolute position, should, after correction for blinking, give three localisations with a probability of 0.68^3 (i.e. p=0.32) that all three will fall within a 20 nm circle of the absolute position (Figure 3.6). Better precision means greater certainty of the location of the fluorophore, which would make clustering easier. Precision can be improved by collecting more photons per fluorophore.

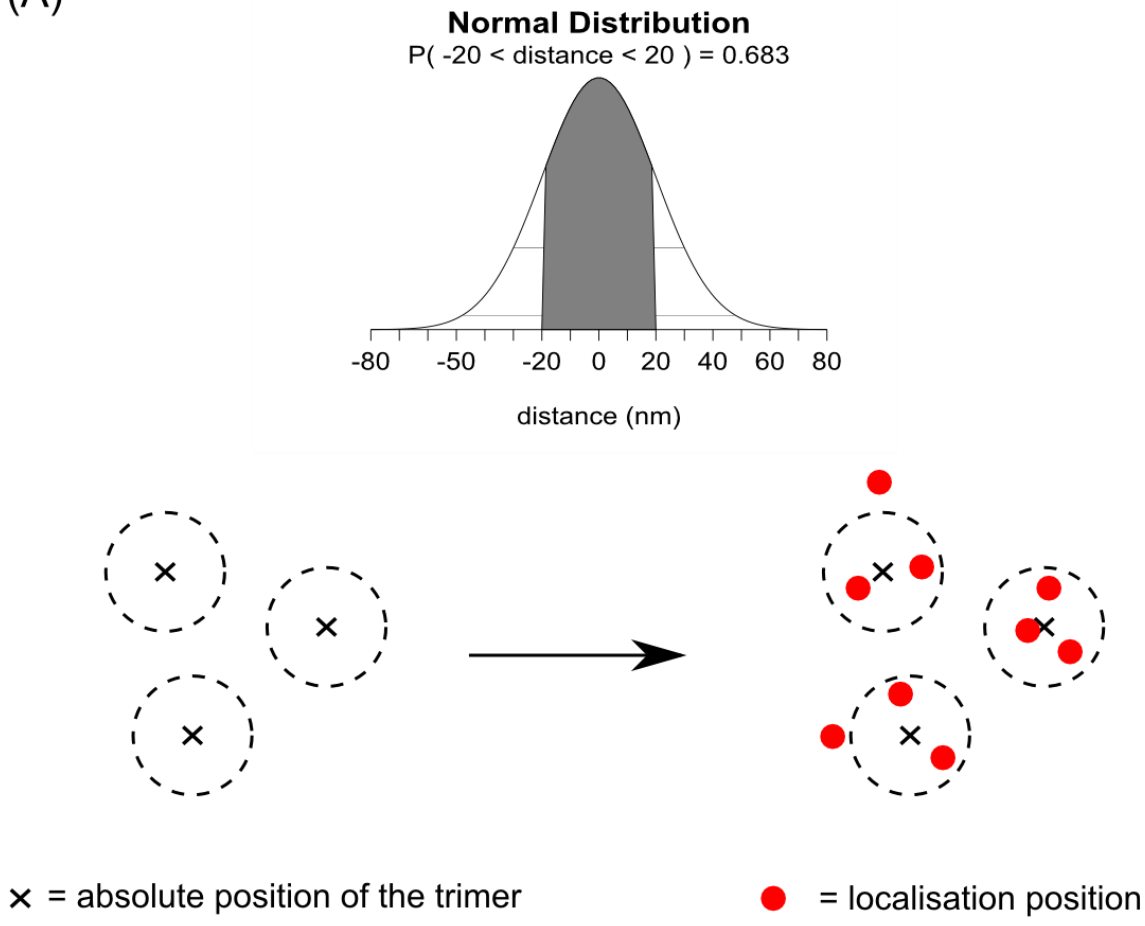
Figure 3.6 Precision is the standard deviation of the probability distribution of seeing a localisation around a molecules absolute position.

There is inherent uncertainty in localising the position of a fluorophore from an image on a camera. The probability distribution of where a localisation will be seen around the absolute position of the fluorophore can be modelled as a normal distribution and the precision, or error in localisation, is generally considered to be one standard deviation around the mean.

(A) A simulated normal distribution with a standard deviation of 20, which is what is expected from the microscope set up used, and a diagram of how localisation position would relate to absolute position. This means that for any given fluorophore position 68% of localisations will be found in a 20 nm radius of the absolute position. Therefore a trimer of mEos3.1, assuming each unit of mEos3.1 has the same absolute position, should, after correction for blinking, give three localisations with a probability of 0.68^3 (i.e. p=0.32) that all three will fall within a 20 nm circle of the absolute position.

(B) The distribution of localisation positions from the Peak Fit fitting software when 5000 spots are simulated in the same absolute position using the parameters of the microscope. The histogram of localisation positions form a normal distribution around the absolute position of the simulated spot in both *x* and *y*. A Q-Q plot can be used to compare two different distributions by plotting the quantiles of each distribution against each other. In this case the quantiles of the localisation position distribution for *x* and for *y* have been plotted against a theoretical normal distribution; the fact that this plot gives a roughly straight line is a measure of its normality.

(A)


Normal Distribution
P(-20 < distance < 20) = 0.683
-80 -50 -20 0 20 40 60 80
distance (nm)
× = absolute position of the trimer
= localisation position
= precision

(B)

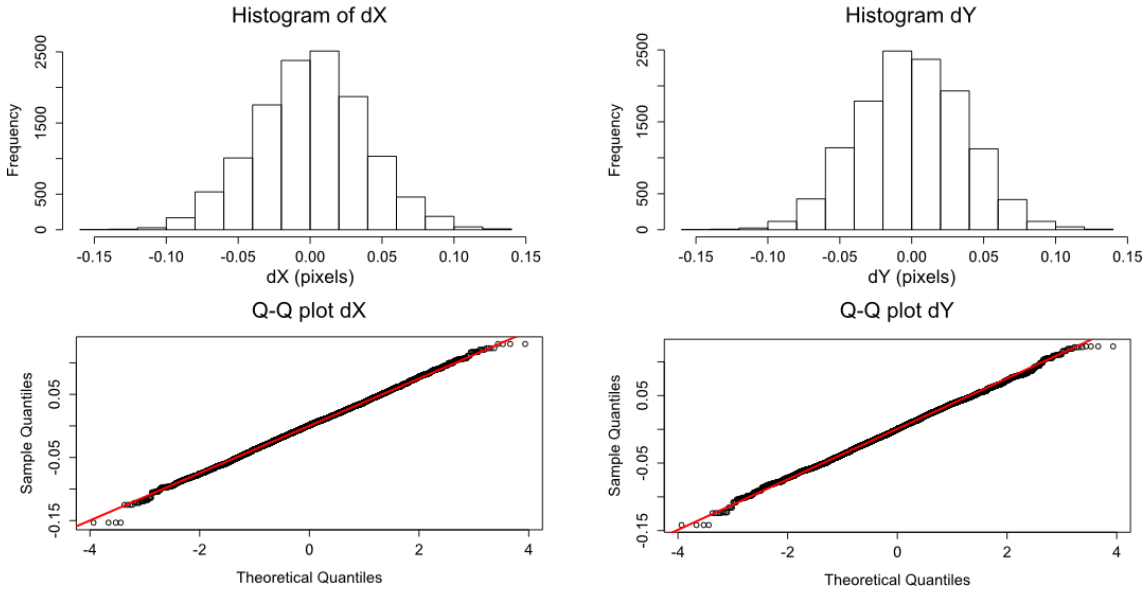

Histogram of dX
Frequency
0 500 1500 2500
-0.15 -0.10 -0.05 0.00 0.05 0.10 0.15
dX (pixels)
Histogram dY
Frequency
0 500 1500 2500
-0.15 -0.10 -0.05 0.00 0.05 0.10 0.15
dY (pixels)
Q-Q plot dX
Sample Quantiles
-0.15 -0.05 0.05
-4 -2 0 2 4
Theoretical Quantiles
Q-Q plot dY
Sample Quantiles
-0.15 -0.05 0.05
-4 -2 0 2 4
Theoretical Quantiles

Photo-activation efficiency is the probability of being able to image a given fluorophore. With PA-FPs a proportion of fluorophores will not be activatable, reasons for this include premature photo-bleaching, the protonation state of the native form (Adam *et al.*, 2009) and potential mis-folding. A table has been published of photo-activation efficiencies for a number of commonly used PA-FPs, produced by using them in conjunction with a protein with a known stoichiometry (Durisic *et al.*, 2014). This table can be taken as a starting guide when performing experiments, however, since the experiments by Durisic *et al.*, were performed in *Xenopus* it is possible different photo-activation efficiencies will be seen in *S. pombe*.

3.2 Results

3.2.1 Simulating the data to test the limits of the algorithm

There are a number of potential limitations on the ability to use PALM to count cytoplasmic clusters; these include the number of subunits in the cluster, the density of clusters, the photo-activation efficiency of the fluorophore used for tagging and the precision of the PALM system. In order to model these limitations Dr Herbert created a plugin to simulate the behaviour of monomers, dimers and trimers of mEos3.1. It worked by assigning a defined number of clusters randomly across a field of view and then placing molecules (subunits) into each cluster according to a binomial distribution (so for a trimer some clusters would have 3 molecules but others would have fewer). Variables were included for N, the number of molecules in the cluster, p, the probability of seeing a given molecule (modelling the photo-activation efficiency) and the density, which was controlled by the total number of clusters plotted (the field of view was of a defined size). Data were simulated with values of between 0.3 and 0.7 used for p, corresponding roughly to the

values reported for most PA-FPs. To increase realism the algorithm also contained a parameter called precision, which affected where it plotted localisations in relation to the absolute *x,y* positions using a Gaussian distribution. The precision of the experimental system used was ~20 nm so the simulations were performed with precisions of between 5 and 40 nm. The simulations were performed in a plane, so did not include out-of-focus fluorophores, and each molecule in the cluster was assumed to have the same absolute position, i.e. no account for the size of the cluster and the distance between molecules within the cluster was made. The output of the simulation includes a histogram of the frequency of cluster sizes created, including a value for 0 molecules per cluster (modelling a cluster with no photo-activatable fluorophores). However when analysing experimental data it is impossible to count the number of clusters with 0 visible molecules. As such the output histogram from the analysis plugin is actually fitted with a zero truncated binomial. Fitting was done using the assumption that the area under the curve of the zero-truncated binomial must add up to $1 - p(0)$, where $p(0)$ is the probability of zero observations from the standard binomial distribution.

Simulations were performed to assess the limits of the clustering algorithm (see 3.1.4). The first task was to assess the clustering distance threshold parameter required for clustering to be effective on localisations of known precision. As part of the simulation a cumulative distribution of the "intra-molecular" distances was constructed and the 95th, 99th and 100th percentile distances calculated. This was done using the particle linkage algorithm as it only allows singles to join a cluster and all the distances can be recorded. The distance is classed as "intra-molecular" if the localisation joins another from the same cluster (i.e. correctly joined to the cluster it was assigned to by the simulation), otherwise it is an "inter-molecular" join (i.e. incorrectly joined to a different cluster).

Analysis was done on the intra-molecular distances to determine the range of clustering

distances that must be used to allow correct clusters to form. From the results it can be seen that the variables N (number of molecules per cluster), p (probability of activation) and precision (error of localisation) all have an effect on the value of the 99th percentile distance. Figure 3.7 A shows that as N increases the 99th percentile distance decreases, presumably because the Gaussian is more densely sampled. Additionally, it can be seen that as p increases the 99th percentile distance decreases (except where N=2) and that the separation between the 99th percentile distances of different precisions is greatest at higher p values (Figure 3.7 B). It can also be seen that the effect of the other two variables is swamped by the effect of precision, with a lower precision giving a lower 99th percentile distance (Figure 3.7 C). The 99th percentile distance was also plotted against Np, which is the number of molecules multiplied by the photo-activation efficiency, in effect the sampling density, and the more densely the cluster was sampled the lower the 99th percentile distance. These simulations suggest that a clustering distance threshold of roughly 3× precision should be used, for example where precision = 20 nm the 99th percentile distance is approximately 60 nm.

Figure 3.7 Graphical representations of the intra-molecule distances in simulated clusters

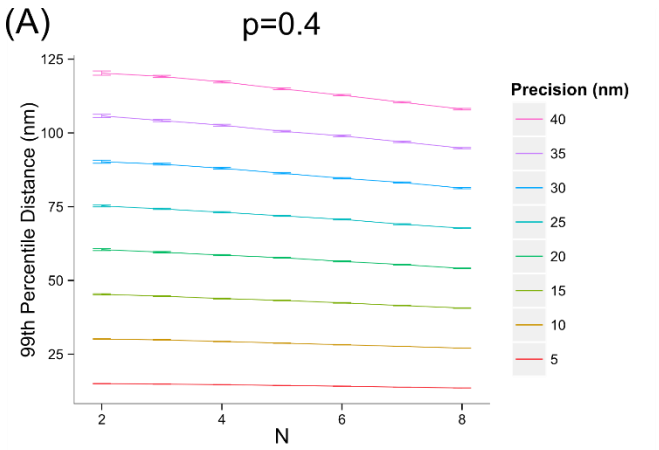
Simulated molecule distributions were produced using different parameters for N (cluster size), p (photo-activation efficiency) and precision. Localisations were assigned to clusters using the particle linkage algorithm as it only allows singles to join a cluster and all the distances can be recorded (see 3.1.4). As part of the simulation a cumulative distribution of the distance between a localisation and the cluster it joined (intra-molecular distance) was constructed and the 95th, 99th and 100th percentile distances calculated.

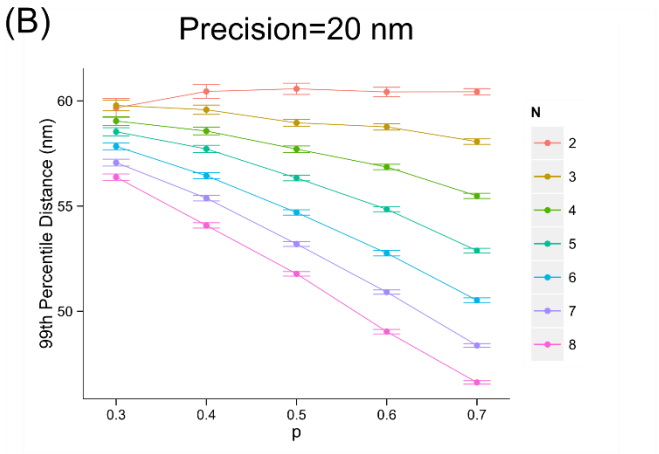
(A) The 99th percentile of distances was plotted against N for a variety of precisions and for a p of 0.4 (expected p of mEos3.1). The 99th percentile of distances value decreases slightly as N increases.

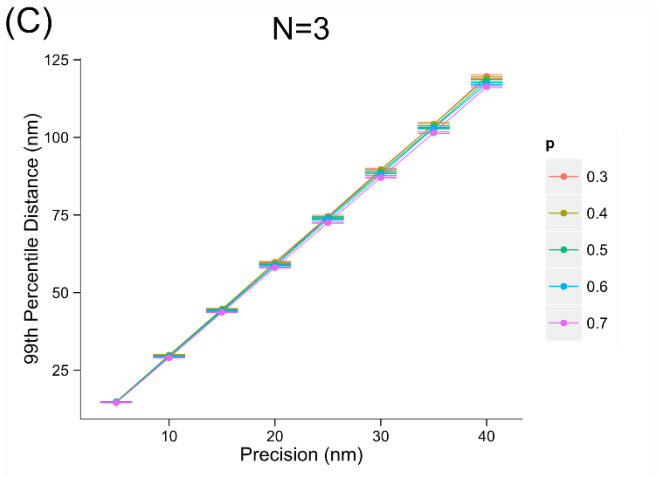
(B) The 99th percentile distance decreases as p increases for a range of N values at a precision of 20 nm, approximately the precision of experiments.

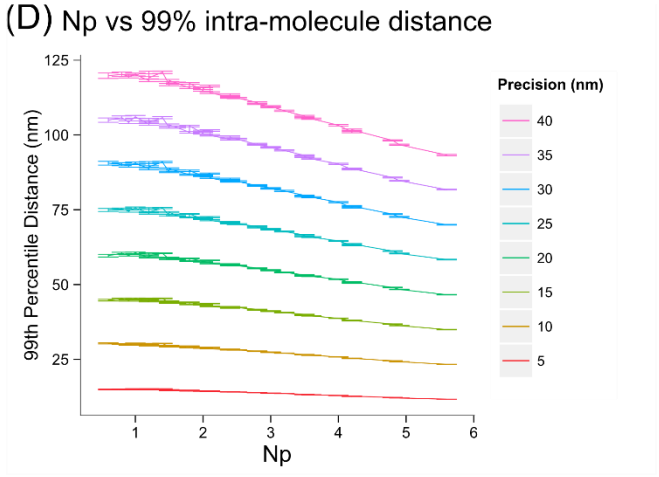
(C) Precision has the strongest effect on the 99th percentile distance, with a higher precision value (greater localisation uncertainty) resulting in a higher 99th percentile value.

(D) The 99th percentile distance plotted against Np which is the number of molecules per cluster multiplied by the photo-activation efficiency, in effect the sampling density. At higher sampling density the 99th percentile distance is reduced as the localisations become more tightly packed.


(A)
p=0.4
99th Percentile Distance (nm)
125
100
75
50
25
2
4
6
8
N
Precision (nm)
40
35
30
25
20
15
10
5


(B)
Precision=20 nm
99th Percentile Distance (nm)
60
55
50
0.3
0.4
0.5
0.6
0.7
p
N
2
3
4
5
6
7
8


(C)
N=3
99th Percentile Distance (nm)
125
100
75
50
25
10
20
30
40
Precision (nm)
p
0.3
0.4
0.5
0.6
0.7


(D) Np vs 99% intra-molecule distance
99th Percentile Distance (nm)
125
100
75
50
25
1
2
3
4
5
6
Np
Precision (nm)
40
35
30
25
20
15
10
5

Another factor with the potential to interfere with the ability of the algorithm to assign clusters correctly was the density of the clusters. Where clusters are at a low density they are unlikely to overlap so there are few instances where the algorithm will join two clusters as one. However where clusters are at a high density it is likely they will overlap and the algorithm will join two or more clusters together giving an over-counting error. Figure 3.8 C shows the results of clustering on low and high density data, where clustering can incorrectly group localisations into a large cluster when the true positions are closer than the clustering distance. Analysis of the intra-molecular distances demonstrated that most clusters (99%) can be correctly assigned using approximately 3× the localisation precision for the clustering threshold.

To investigate the effect of cluster density on the clustering algorithm, data were simulated using a range of densities, N, p and precision values and clustered using 3× the precision as the clustering threshold. This allowed calculation of the density limit where the clustering analysis produced the correct N and p values. Clustering was considered correct when the algorithm returned the correct N value and a p value within 10% of the simulated value. This analysis was performed twice: once for an unknown (variable) N where the binomial can fit to any N, e.g. when the stoichiometry of the cluster is not known; and once for a known (limited) N where the binomial is limited to a maximum N, e.g. when the stoichiometry is known from complimentary experiments (see Appendix 8.5.1 for the full results from limited N fitting, 8.5.2 for variable N fitting). The simulations show that the density limit for variable N is lower than for a limited N. The precision has a very strong effect on the densities at which clustering is successful, the smaller the localisation precision value the higher the density that can be correctly analysed. The p value also has an effect on the density that can be analysed, with a greater p value resulting in a greater maximum density at which N is correctly returned (Figure 3.8). With a photo-activation efficiency of 70% the density of clusters can be higher before N is incorrectly fitted than if

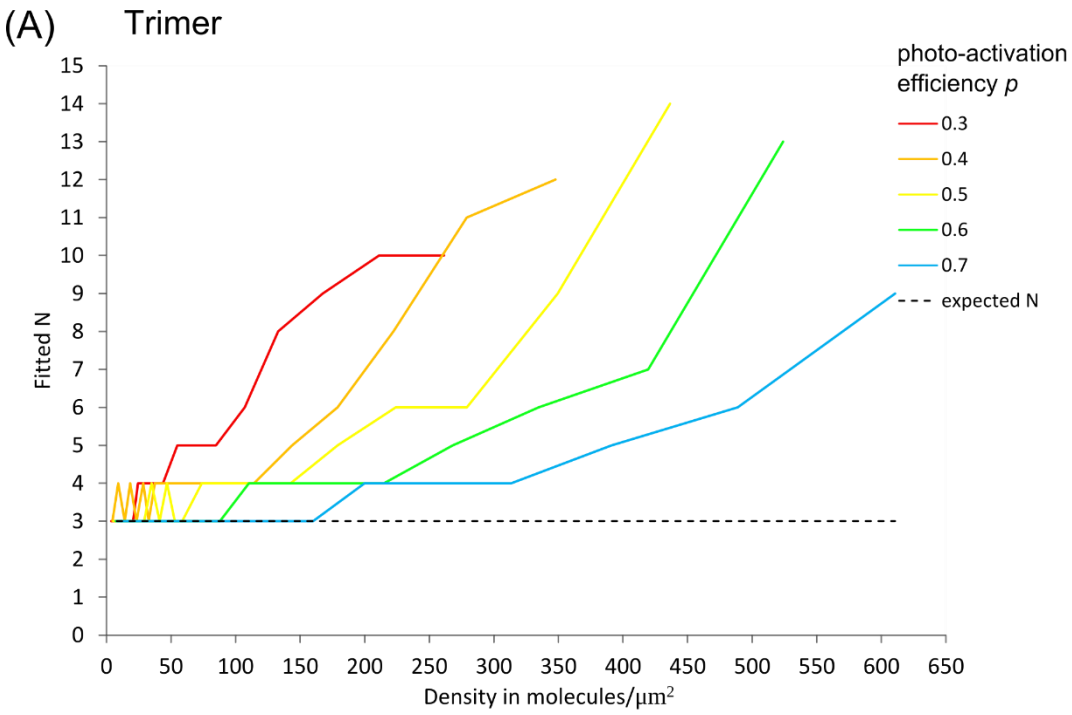
the photo-activation efficiency is 40%, which is what is expected for mEos3.1. This means that the choice of PA-FP is very important. The *N* value has a slight effect, as *N* increases so do the densities that can be analysed. It should be noted that this is the density of clusters, where one cluster could be a monomer, a dimer or a trimer (as opposed to the localisation density where 1 trimer = 3 localisations). As shown above, as *Np* increases the intra-molecular distance needed for clustering decreases. This is because higher *Np* effectively means a greater sampling frequency, greater sampling frequency will cause each cluster to appear more closely packed and therefore easier to group together. This means that trimers should be easier to analyse *in vivo* and explains why simulated trimers can be correctly clustered at higher densities than dimers.

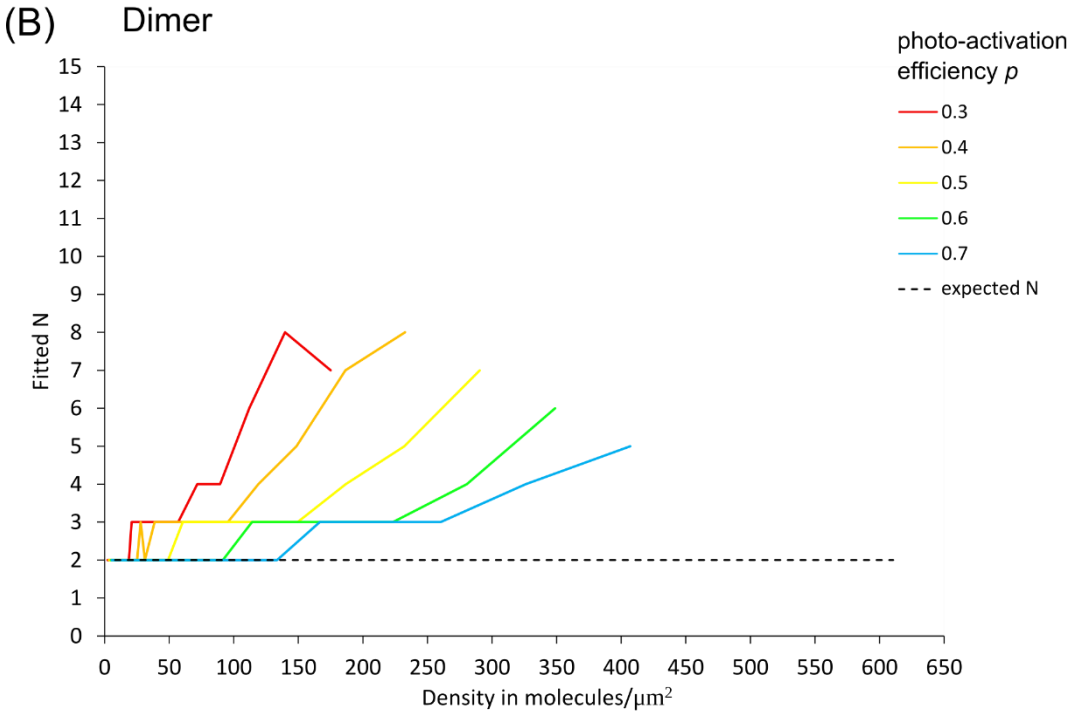
Figure 3.8 The ability of the algorithm to correctly fit *N* varies with molecular density and photo-activation efficiency for a trimer

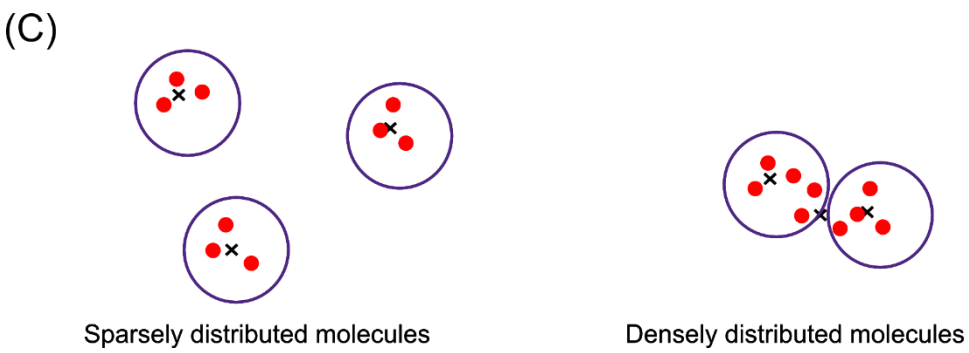
(A) Predicted *N* for mEos3.1 trimer is shown as a function of molecule density for various photo-activation efficiencies, *p*. The true *N* (3) is shown as a dashed line. Higher densities result in over-estimation of N.

(B) The same for simulations of mEos3.1 dimer with a true *N* of 2.

(C) The density of the oligomers can lead to overestimates of *N*. The black x shows the absolute position of a hypothetical trimer, the red dots show the localisation positions for each subunit of the trimer and the blue circles indicate the how the algorithm would cluster them. Where the oligomers are sparse there is no overlap between the areas of precision and it is easy to cluster the localisations correctly, where the oligomers are dense the areas of precision overlap and it is possible to incorrectly assign clusters. In this example three trimers are correctly clustered in the sparsely distributed image (left) but they are incorrectly clustered in the densely distributed image (right).


(A)
Trimer
photo-activation efficiency p
0.3
0.4
0.5
0.6
0.7
expected N
Fitted N
Density in molecules/µm²


(B)
Dimer
photo-activation efficiency p
0.3
0.4
0.5
0.6
0.7
expected N
Fitted N
Density in molecules/µm²


(C)
Sparsely distributed molecules
Densely distributed molecules

The simulations show that at a localisation precision of 20 nm and a photo-activation efficiency for mEos3.1 of 0.4 in idealised circumstances (i.e. perfect fluorophore identification) trimers can be analysed at densities of <9 molecules/μm^2 for variable fitting and <32 molecules/μm^2 for limited fitting and dimers at densities of <6 molecules/μm^2 for variable fitting and <12 molecules/μm^2 for limited fitting (see Appendix 8.5.3 for full table of results). For biological questions it is more likely that the variable fitting will be used. In the case of the artificial mEos3.1 oligomers, the single units in the dimer and trimer constructs are separated by 12 amino acids, estimated to give about 5 nm length, which would, in effect, decrease the localisation precision and reduce the density that can be analysed. In addition an *S. pombe* cell is usually 3-4 μm in diameter and the depth of focus in this system is 500 nm meaning that a number of the fluorophores will be out-of-focus and therefore have a much higher uncertainty or else contribute to the background noise of the image. This means that for real data *in vivo* the workable densities could be even lower. However, since the shift out-of-focus is continuous, if only localisations with similar PSF widths (meaning they are in the same region of focus) are clustered this effect can be mitigated. The implication of this density limit on biological data is considered further in the discussion section (see 3.3.1).

3.2.2 Comparing the simulated data to experimental data in cells

In order to establish how much more limited quantitative PALM was *in vivo* compared to *in silico* it was necessary to compare the simulations to experimental data using the monomer, dimer and trimer constructs in cells. PALM imaging data were collected using cells expressing the trimer *spo5* construct. Cells were grown to roughly mid-log phase un-induced and fixed before imaging with continuous low levels of 405 nm activation light

during acquisition. Localisations were fitted using the Peak Fit ImageJ plugin (written by Dr Herbert) and re-plotted on a 10× scale using the width of the PSF (Figure 3.9). Cells were manually selected for analysis using the fluorophore distribution throughout the image sequence; localisations should be easy to identify above the background and the distribution of observations consistent through the time course. A selection of cells was manually outlined and the sub-set of localisations within the cells used for clustering analysis (Figure 3.9). The area of the selection was measured and used to calculate the density of localisations after tracing. Using the clustering analysis described by Puchner *et al.*, it was expected that N=1, N=2 and N=3 would be seen for the monomer, dimer and trimer respectively and p=0.4 would be seen for dimer and trimer (N=1 always returns a p=1) for cells in which the localisation density was below the threshold suggested by the simulation analysis (see section 3.2.1).

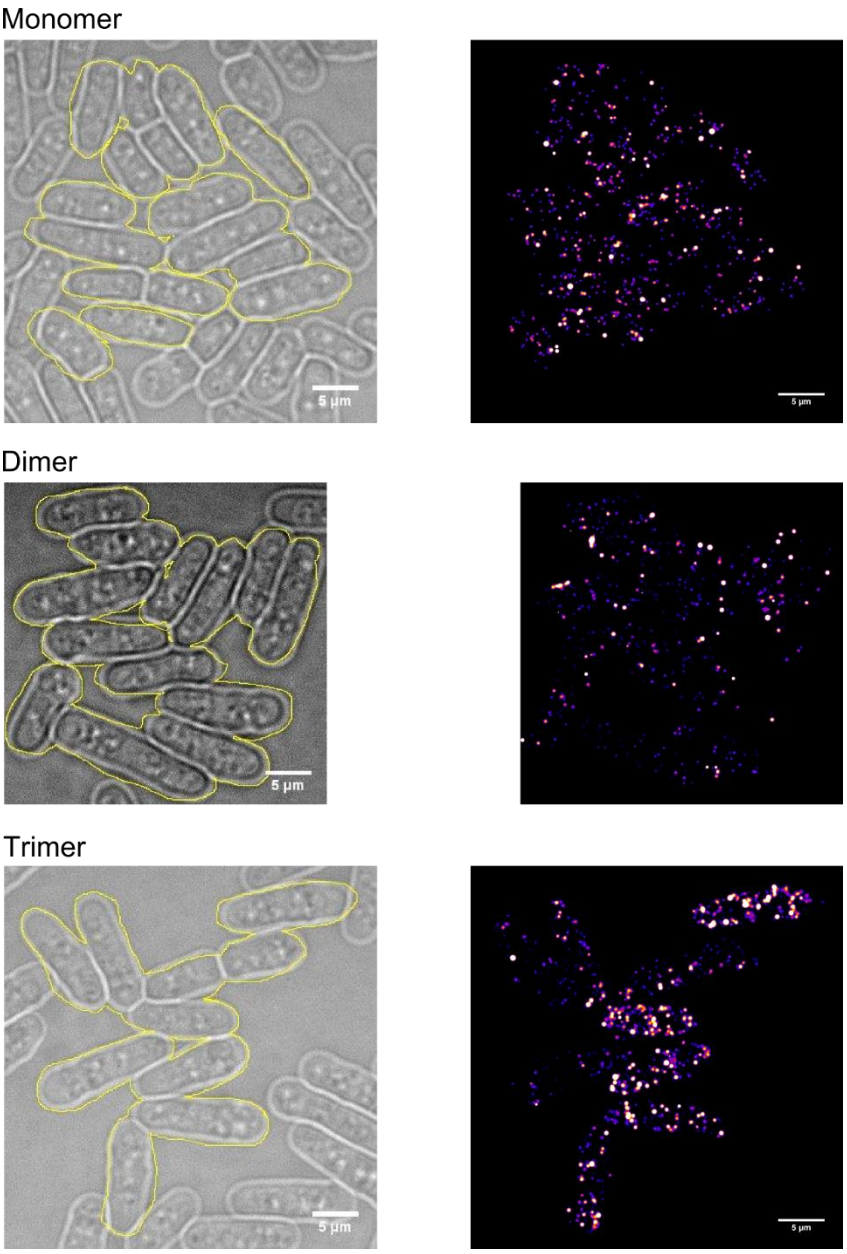


Figure 3.9 The white light and localisation images from a monomer, dimer and trimer dataset

White light and localisation images of cells expressing the monomer, dimer and trimer mEos3.1 constructs. The morphology of the cells is normal, suggesting that the constructs are not interfering with cell growth. The cells do not show the morphology of stationary phase cells (shorter and rounder). The yellow outline indicates the area of the mask used for filtering the results. Localisation images are plotted with the width of the PSF for each localisation.

3.2.2.1 Calculating the tracing parameters

Before the clustering analysis could be run it was necessary to trace the data. Tracing the data allows multiple localisations that appear close together in both time and space to be combined into a single localisation as they most likely come from a single fluorophore. This eliminates over-counting resulting from blinking (Annibale *et al.*, 2011) or from a single fluorophore being active across more than one frame. In the Puchner *et al*, (2013) paper a combination of the spatial density analysis and a dark time analysis were used to determine the time and distance thresholds they would use for tracing. In the spatial analysis they noted that the majority of increased density was below $g(r) = 150$ nm so they used this value as their distance parameter for an analysis of the dark times in their data. The dark time is the time between localisations in a defined space (in their case within 150 nm of each other). By plotting a cumulative histogram of dark times they calculated that 99% of their data came within 2.66 seconds, which they used as the time threshold for tracing.

When the density analysis was performed on the data collected on monomer, dimer and trimer of mEos3.1 it could be seen that the majority of the density for this system falls below 200 nm (Figure 3.10).

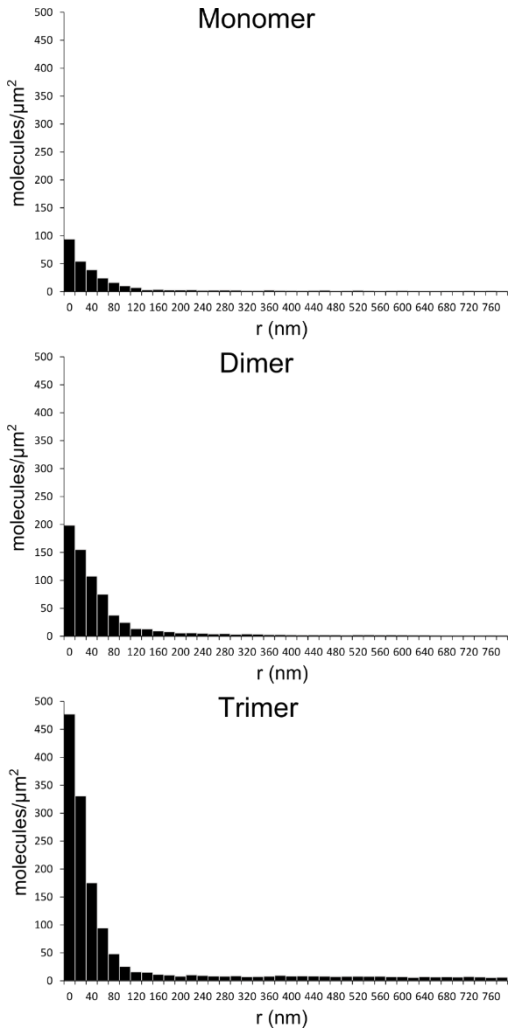


Figure 3.10 Untraced spatial distribution analysis of cells expressing a monomer, dimer or trimer of mEos3.1

The spatial distribution analysis gives a measure of the density of molecules at a variety of distances out from any given molecule. The number of molecules/μm^2 was plotted in a histogram against the distance (in nm) with a bin width of 20 nm up to a maximum distance of 800 nm. This was repeated for data taken using cells expressing the monomer, dimer and trimer constructs of mEos3.1. Peaks at short distances in the histogram show that the localisations are clustering. Some of the clustering seen is likely to be due to blinking of the same molecule and thus can be removed by tracing localisations through time. It can be seen that the majority of increased density is below 200 nm for these analyses.

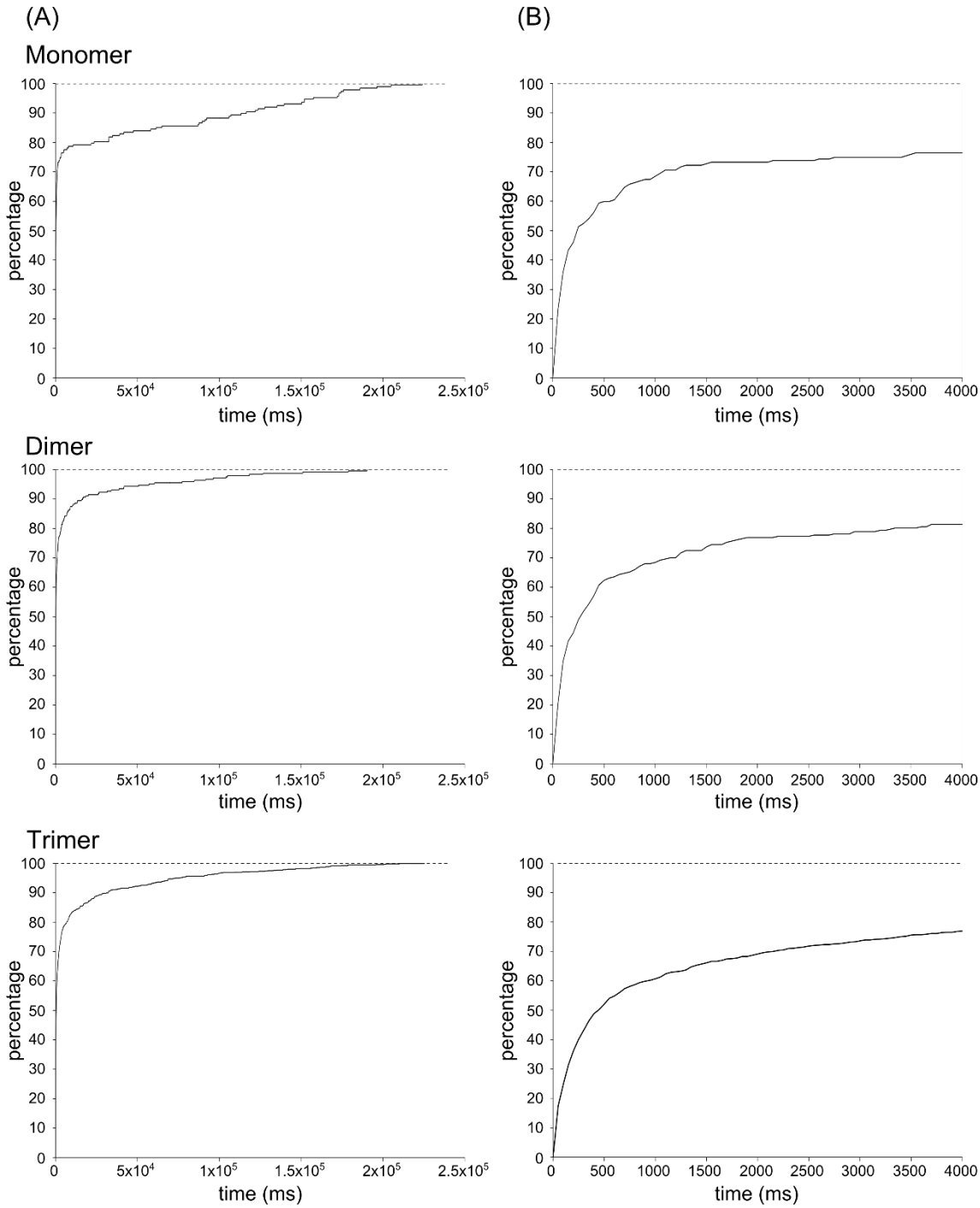
The dark time analysis used by Puchner *et al.*, was repeated for data collected from cells expressing a monomer, dimer and trimer of mEos3.1. The parameter of 200 nm as suggested by the spatial distribution analysis was used for this dark time analysis. The cumulative histograms show a small shoulder followed by a long tail with a constant slope (Figure 3.11). The cumulative histograms reach 100% much more slowly than the ones shown by Puchner *et al.*, and the 99th percentile is much higher than 2.66 seconds for all of the constructs (204 s for monomer, 151 s for dimer and 168 s for trimer). This could be because a much longer-lived dark state than observed by Puchner *et al*, was being seen. However it is more likely that the long shallow curve is due to noise, which can occur all over the image and at any time. Puchner *et al.*, were working on membrane localised proteins which may have enabled them to exclude background from out of focus fluorescence more easily. If a lot of noise that is not mEos is being fitted it may be possible to change the parameters to get rid of the noise, the problem being that this may affect the recall rate and lead to throwing away of correct localisations with false positive noise-based localisations.

Figure 3.11 Dark time analysis for cells expressing a monomer, dimer and trimer of mEos3.1

A cumulative histogram of dark times for data obtained using cells expressing a monomer, dimer and trimer of mEos3.1 in the cytoplasm. Puchner *et al.,* used the 99th percentile from a similar analysis as their time threshold for tracing. Dark times were analysed with a distance parameter of 200 nm obtained from spatial distribution analysis and a max dark time of 800 s.

(A) Cumulative histograms of dark times were plotted for each of the monomer, dimer and trimer constructs and the 99th percentile calculated (204 s for monomer, 151 s for dimer and 168 s for trimer). Unlike the histograms shown by Puchner *et al.*, this data reaches 100% slowly with a long shallow tail. This is possibly a result of fitting noise.

(B) The same data replotted without the upper portion of the curve attributed to noise. As such the time at the shoulder of the curve as the gradient drops significantly was considered as an upper limit for tracing localisation data from the same molecule. Looking closer at this part of the histogram suggested a time of roughly 1 s.


(A)
(B)
Monomer
Dimer
Trimer
percentage
time (ms)
0
10
20
30
40
50
60
70
80
90
100
5x10^4
1x10^5
1.5x10^5
2x10^5
2.5x10^5
500
1000
1500
2000
2500
3000
3500
4000

In the spatial analysis of the monomer, dimer and trimer constructs (Figure 3.10) the majority of the data fall below 200 nm, as such manual tracing was initially performed with a distance threshold of 200 nm. The dark time analysis gave a 99th percentile time threshold of 204 s, 151 s and 168 s for monomer, dimer and trimer respectively which was used as the time threshold for manual tracing. Initially these numbers were used for tracing and the spatial analysis repeated. If the blinking had been successfully corrected it was expected that the distribution for monomer would be flattened whereas for the dimer and trimer a peak would continue to be seen. However with these parameters all of the distributions were flattened suggesting over-tracing (Figure 3.12 A). The time parameter used by Puchner *et al.*, at 2.66 s was much lower than the 99th percentile found by dark time analysis of this data. The upper, approximately linear, section of the cumulative dark time histograms is likely due to noise. Assuming this can be ignored, then the parameter for the clustering analysis can be taken from the shoulder of the histogram where the gradient rapidly drops (i.e. the majority of true blinking events have now been counted). Looking at the shoulder of the cumulative dark time histogram suggested a time parameter of 1 s. Spatial distribution analysis of localisations traced using 200 nm and 1 s shows that the dimer and trimer retain their clustering peak, however with this time parameter the monomer distribution also shows a small clustering peak at short distances, suggesting the blinks have not all been traced (Figure 3.12 B). Finally the time parameter was optimised manually until the spatial analysis on the monomer showed a flat distribution. A time parameter of 10 s was able to flatten the monomer distribution but leave a peak for the dimer and trimer distributions (Figure 3.12 C).

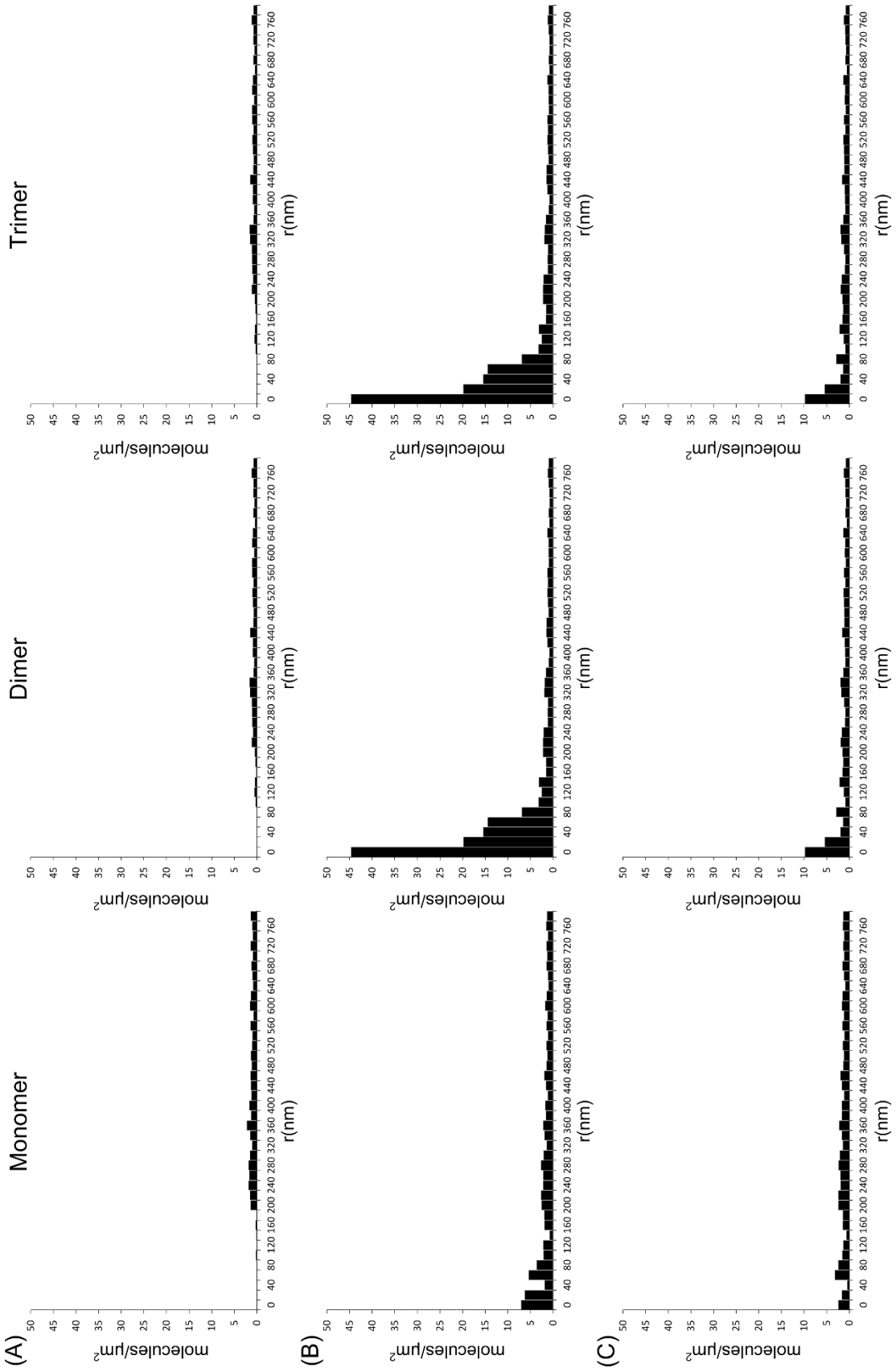
Figure 3.12 Spatial distribution analysis of tracing parameters

Spatial distribution analysis was repeated on data that had been traced using the distance threshold of 200 nm suggested by the original spatial distribution analysis (Figure 3.10) and a variety of different time parameters.

(A) Dark time analysis (Figure 3.11 A) gave a 99th percentile time of 204 seconds for the monomer, 151 seconds for the dimer and 168 seconds for the trimer. The data were traced with a distance parameter of 200 nm and the relevant time parameter. It can be seen that this eliminates the peak at short distances and flattens the histogram for all of the constructs suggesting that these parameters have led to over-tracing.

(B) A time of 1 second, representing the shoulder of the cumulative dark time histogram (Figure 3.11 B), was used to trace the data. It can be seen there is still clustering at short distances for all constructs (note that the Y scales are not equal).

(C) Tracing was repeated using a time of 10 s that was found to flatten the monomer distribution without flattening the dimer and trimer distributions.


(A)
Monomer
Dimer
Trimer
(B)
(C)
molecules/μm²
r(nm)
0 40 80 120 160 200 240 280 320 360 400 440 480 520 560 600 640 680 720 760
0 5 10 15 20 25 30 35 40 45 50

3.2.2.2 Clustering after tracing does not give correct *N* and *p* values

Clustering analysis was performed using all the tracing parameters described in 3.2.2.1 in cells with an expression level such that the density should be within the limits suggested by the simulations (3.2.1). The values for *N*, *p* and density were recorded for all of the tracing parameters (Table 3-1). As predicted by the spatial distribution analysis the time threshold suggested by the 99th percentile of the cumulative dark time over-traces leading to only monomers. The time threshold of 1 s did not eliminate the clustering peak at short times in the spatial distribution analysis and indeed led to over-estimates of the *N* value in clustering. The time threshold of 10 s appeared to flatten the monomer spatial distribution whilst leaving peaks in the dimer and trimer distributions, however it still led to over-counting in the monomer and dimer, and although it returned the correct *N* value for trimer it returned a very low *p* value. With the exception of the over-tracing where *p* is returned as 1 all the *p* values are much lower than the expected 0.4 (Durisic *et al.*, 2014) which casts doubt on whether the counting analysis can be trusted even where it appears to return the correct *N* value.

Expected		Distance threshold (nm)	Time threshold (s)	Density	Calculated	
N	***p***				***N***	***p***
1	0.4	200	204	0.931	1	1.000
2	0.4	200	151	0.636	1	1.000
3	0.4	200	168	1.579	1	1.000
1	0.4	200	1	1.020	2	0.012
2	0.4	200	1	0.729	4	0.026
3	0.4	200	1	2.135	5	0.044
1	0.4	200	10	1.051	2	0.037
2	0.4	200	10	0.691	4	0.005
3	0.4	200	10	1.937	3	0.031

Table 3-1 Results of clustering algorithm

The data from cells expressing monomer, dimer and trimer was traced using three different parameter sets and the *N*, *p* and density recorded for each.

There are a number of reasons the *p* values may be so low: The first is that because continuous activation rather than pulsed activation was used, overlapping fluorophores may have been being activated in the same frame. This would give rise to either a single localisation or a localisation with a broad PSF which was discarded as out-of-focus. A second possibility is that not all of the available fluorophores have been activated. In PALM experiments the likelihood of activating a fluorophore is highest at the beginning of the experiment when few have been photo-bleached and this gradually decreases over the course of the experiment meaning that the number of activations follows an exponential decay over time. However the experiments that produced this data were performed under the imaging regime suggested by Lee *et al.*, (2012) which uses continuously modulated 405 nm power until no further activations were seen by the end of acquisition. Using this approach should have mitigated the possibility of un-activated fluorophores. A further possibility is that the trimer constructs are not folding in a way that allows all three of them to be fluorescently active. Or it could be that conditions in *S. pombe* are not conducive to proper fluorophore maturation and so the photo-activation efficiency seen is actually lower than that reported by Durisic *et al.*, (2014). Assuming the *p* values calculated are true, then the probability of seeing a given molecule tagged with mEos3.1 is much lower in the system used than the probability reported by Durisic *et al.*, meaning that it is unlikely to be possible to see all of the molecules in a complex. Alternatively the *p* values calculated are not correct and there may be a problem with the localisation or clustering algorithms.

3.2.3 Replicating the simulations using a lysate

Having been unable to consistently achieve a correct value for *N* and *p* in cellular data an attempt was made to perform the experiment using a lysate spread onto a slide so that the density could be controlled. This would help establish whether the mEos3.1 constructs

could be counted at all. A protocol was established for lysing the cells using glass micro-beads and a ribolyser (see Materials and Methods 2.3.4) and the proteins encouraged to stick to the slide using Poly-L Lysine (Sigma). The main difficulty with the lysate approach was that there was nothing on the slide to use as a focus guide, making it hard to tell whether the focus was on the lysate or on dust or other contaminants on the slide. This was mitigated by attempting to achieve focus in the first few frames of the acquisition once activation had been started.

The slides were cleaned thoroughly for several hours in an ozonator with a UV light source. However, there was still fluorescence that could be seen that was bright but non-activatable and appeared to bleach more slowly than mEos3.1, adding unwanted background. Pulsed activation (see Materials and Methods: 2.3.4.1) was used in order to differentiate activated fluorophores from background fluorescence. Spots were seen that appeared to be activating after a pulse of 405 nm light (Figure 3.13 A). When the number of localisation per frame was plotted it could be seen that, in a sample of cells expressing mEos3.1 trimer, the number of localisations peaks at the time of the 405 nm pulse (Figure 3.13 B), whereas in the trimer lysate sample (Figure 3.13 A) there were a large number of localisations early on in the acquisition but the peak in localisation number following the 405 nm pulse was lost in the noise. This demonstrates that, in contrast to what is seen in cells, pulsed activation is not photo-converting fluorophores in the lysate in sufficient numbers to be visible over background. A lysate was also produced from wild type AW310 cells, cells which should not contain any mEos at all, and in the reconstructed image, even after tracing and filtering, a similar number of localisations can be seen to those from the trimer lysate (Figure 3.13 C). Thus the mEos3.1 is contributing a negligible amount of localisations to the lysate total, or none at all. The wild type lysate shows an increase in the number of localisation over time, before falling off. This is not observed in the mEos3.1 lysate. This may be because of an exceptionally high density of fluorophores (due to non-

uniform spreading) resulting in failure to fit localisations (the fit procedure is sensitive to overlapping fluorophores in the same image frame). As some of the auto-fluorescent molecules were photo-bleached it became more possible for the algorithm to localise what remained. It may also be worth noting that in fixed trimer cells there were a lot of localisations found in the membranes that are unlikely to correspond to the mEos3.1 trimers (which should be cytoplasmic). When the cells are broken open these fluorescent molecules from the membrane will be spread onto the slide alongside any mEos3.1. Given the difficulties in being certain that there is any mEos3.1 present it was not probed further as a technique.

Figure 3.13 Comparison of super-resolution imaging of a cell lysate with whole cells

Super-resolution image data were collected using a cell lysate of AW 310, a wild type strain with no mEos, fixed cells expressing a trimer of mEos3.1 and a lysate of mEos3.1 trimer. A histogram of the localisations per frame was plotted (Note the *y* scale varies). The purple lines indicate the time at which a pulse of 405 nm activating light was applied. Localisation number increases following a pulse should be due to mEos3.1 activation (photo-conversion). Super-resolution localisations were plotted on a 10× scale as two-dimensional Gaussians with a width equal to the precision.

(A) A lysate made from cells expressing the trimer construct. The localisations plotted do not peak with the 405 pulses. Localisations are seen across the field of view.

(B) A lysate made from wild type cells with no mEos.

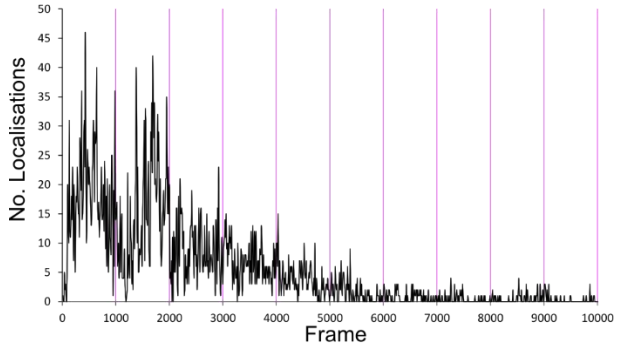
(C) Cells expressing the trimer construct. Localisation numbers peak with the 405 pulses and localisations are limited to the inside of cells.

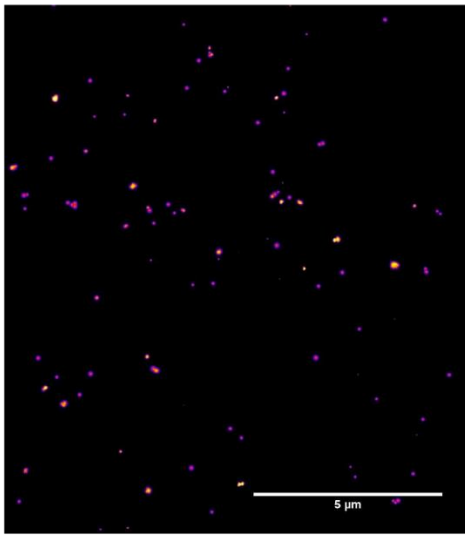
Histogram of localisation numbers

Super-resolution image

(A)

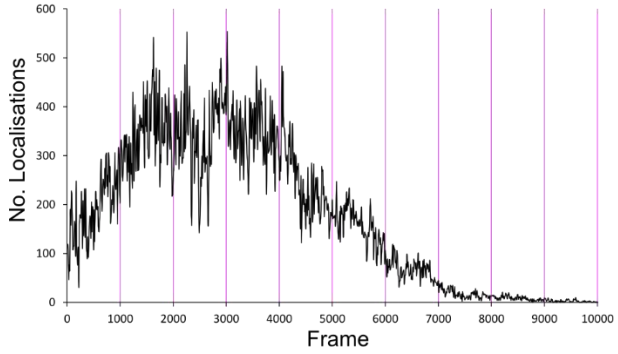
Trimer Lysate

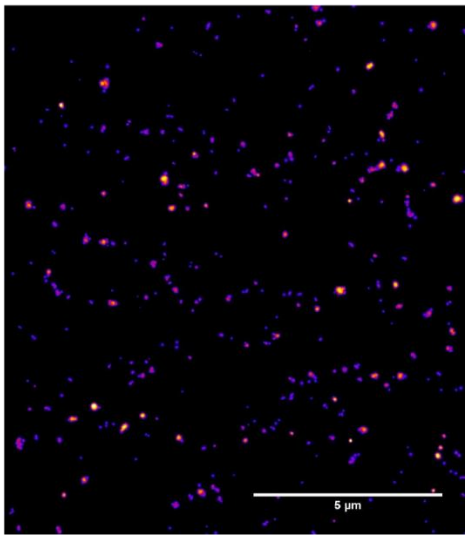




(B)

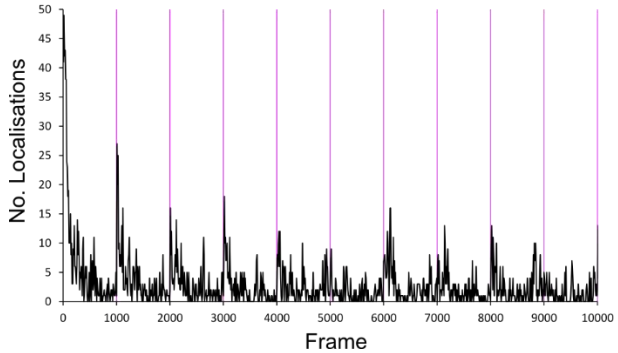
310 (wild type) lysate

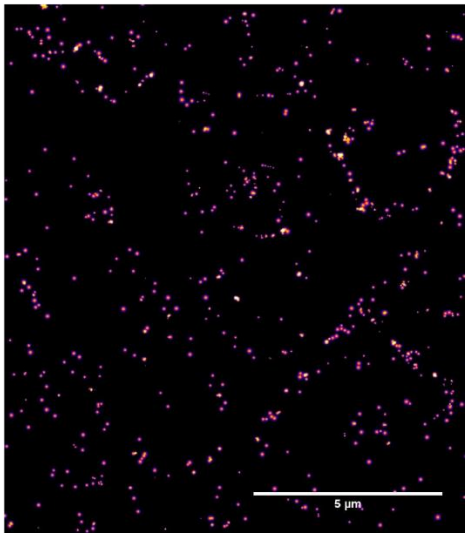




(C)

Trimer in cells





3.2.4 Counting fluorophores with pulsed activation

The PA-FP mEos3.1 is activated using 405 nm wavelength light, the experiments above used this activation light continuously at very low powers to activate all the fluorophores. However if the 405 nm wavelength light is instead used in a short pulse it gives increased control over activation and photo-bleaching of individual fluorophores. One explanation for the lack of counting success is that the analysis method is sound but that there is a problem with the trimer constructs resulting in there not being sufficient instances where there are three connected, mature and photo-activatable fluorophores. This can be investigated using pulsed activation because the length of time between pulses can be adjusted to ensure all activated fluorophores have been photo-bleached before the next activation pulse. In pulsed activation experiments the localisations can be plotted as a histogram of localisation frequency over frame number. If these histograms show localisation numbers peaking after the pulse and dropping to zero before the next pulse then it can be confidently assumed that all of the blinking events of the activated molecules have been captured. If this is the case the clustering algorithm should be run using only localisations within the interval between each pulse to "cluster" together the blinking events into a single localisation to significantly reduce over-counting from blinking. If the clustering algorithm is run a second time without a time component and the cluster sizes from this second run are a majority of monomers then it strongly suggests that all three mEos3.1 molecules of trimer are not being seen.

This pulsed analysis was performed on cells expressing the mEos3.1 trimer construct. From the histogram of localisation frequency over time it can be seen that the localisations increase after each pulse and drop back down into the noise in the remainder of the cycle (Figure 3.14 A). With the last pulse there is no peak above the noise so it can be assumed that the majority of the fluorescence has been captured. When the clustering analysis was

run for the second time the resultant histogram showed only clusters of size 1 (Figure 3.14 B). This could be because the photo-activation efficiency of the fluorophores is much lower in the *S. pombe* model organism than that seen by Durisic *et al.* (2014). Reasons for reduced photo-activation efficiency include premature photo-bleaching and the protonation state of the native form (Adam *et al.*, 2009). Alternatively it may be that all three mEos3.1 units are unable to properly fold or mature in the string-like construct. Another possibility is that cleavage is separating the mEos3.1 units into monomers (Figure 3.14 C). In the western blot of the mEos3.1 constructs (Figure 3.4) there are bands in the trimer and dimer construct lanes at the same size as the monomer which may correspond to a cleavage product. These results indicate that the synthetic constructs have failed to perform as expected, probably due to *in vivo* malfunction. A better experiment than using synthetic constructs would be to find an endogenous protein complex with a well characterised stoichiometry which is amenable to being tagged with mEos3.1 and use that for testing the counting techniques.

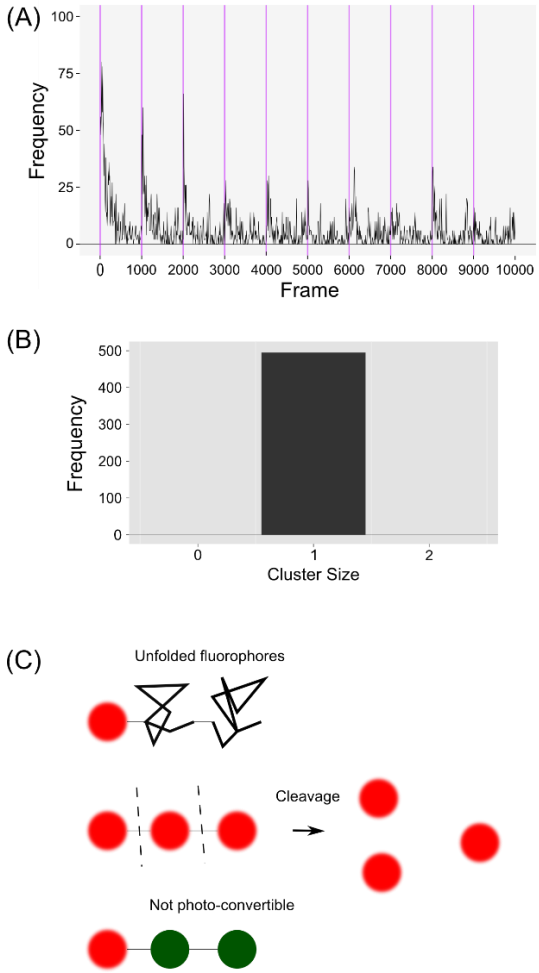


Figure 3.14 Counting fluorophores with pulsed activation

(A) An experiment was performed using cells expressing the mEos3.1 trimer construct and pulsed activation with the 405 nm laser. Localisations were fitted using the Peak Fit plugin and a histogram of localisation frequency over frame number was produced. In a pulsed PALM experiment if a suitable number of pulses and imaging cycles are used, peaks above the noise will cease to be seen after a pulse once all the fluorophores have been imaged.

(B) The clustering algorithm was used with the localisation data but restricted to clustering localisations within the time period between 405 nm pulses in order to group repeat localisations of the same molecule. The clustering algorithm was then run a second time with no time threshold. A histogram was plotted of cluster size against frequency.

(C) Graphical illustration of invalid trimer constructs. Misfolded fluorescent proteins in the trimer, cleavage of the trimer into monomers and non-photo-convertible fluorophores in the trimer.

3.2.5 Creating monomer, dimer and trimer constructs using mEos2

The simulation results showed that the photo-activation efficiency has a moderate effect on the density at which molecules can be clustered; lower efficiency reduces the density that can be correctly analysed. The Durisic *et al*, paper gave a photo-activation efficiency value of 40% to mEos3 but 60% to mEos2. As such, it was hypothesised that the algorithm would be more likely to return the correct value for *N* from constructs of mEos2 than from constructs of mEos3.1. The monomer, dimer and trimer constructs were re-designed to contain mEos2. This involved taking Dr Watson's three mEos3.1 sequences, which were codon optimised for *S. pombe* and changing the codons described by (Zhang *et al*., 2012) to create mEos3.1 from mEos2 (see Introduction 1.2.6), thus creating an *S. pombe* specific mEos2. Three different sequences for mEos2 were designed (A, B and C) in order to reduce recombination within the constructs and these were ordered from a company, Genscript, who supplied the sequences purified in the plasmid pUC57. Sequence optimisation A was cloned into the pAW8ENdeI plasmid (which contains the loxM and loxP sites for cassette exchange (see Materials and Methods: 2.2.4) into the expression locus, P_{urg1}) to give the monomer. Sequence optimisation B was sub-cloned in upstream of A to give the dimer and the sequence optimisation C sub-cloned in upstream of that to give the trimer (Figure 3.15). Each construct would ultimately be expressed under the control of the *urg1* uracil inducible promoter and contained the *spo5* DSR element.

All three plasmid constructs were amplified in *E. coli* and purified using a Machery Nagel midiprep kit and then checked by sequencing. Selected plasmids were transformed into the *S. pombe* strain AW459, which is leucine deficient, and harbours the relevant loxM – loxP expression locus. Colonies were selected for on leucine free EMM plates. Selected colonies were initially checked on the microscope for mEos2 activation and then by western blot. A

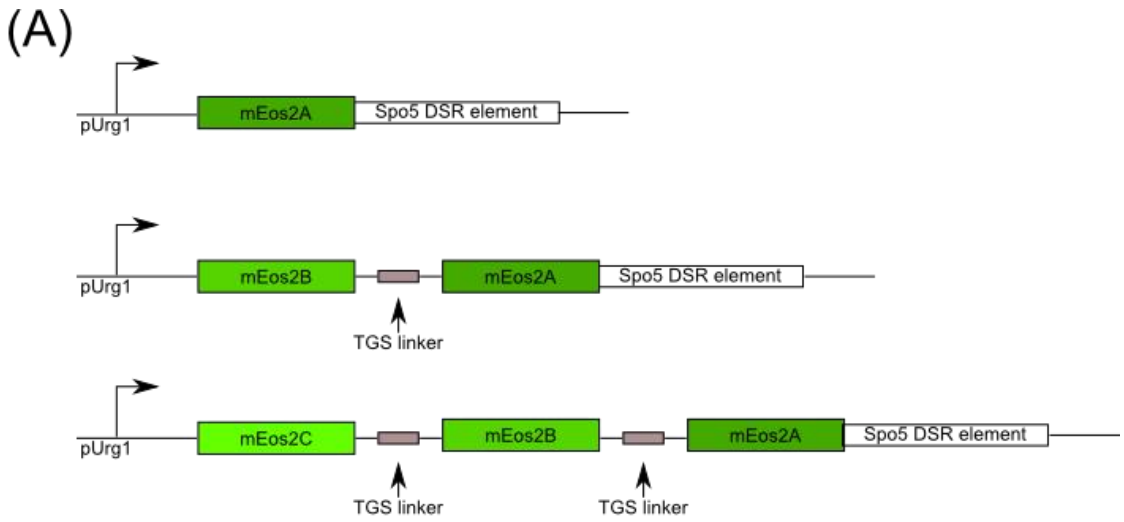
simple TCA extraction (see Materials and Methods 2.2.7) was used and the proteins run on an 8% SDS PAGE gel (see Materials and Methods 2.2.8). A rabbit anti-mEos primary antibody and a swine anti-rabbit IgG horse radish peroxidase (HRP) conjugated antibody were used for visualisation of the proteins. Bands were seen on the western blot corresponding to the projected protein sizes of 28.45 kDa (monomer), 53.77 kDa (dimer) and 80 kDa (trimer) (Figure 3.15).

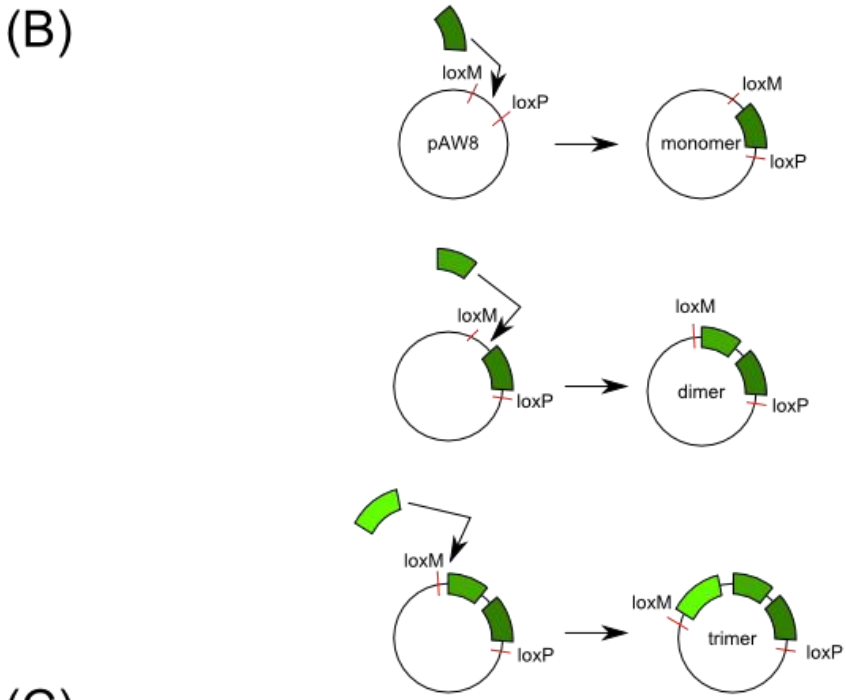
Figure 3.15 Creating a monomer, dimer and trimer of mEos2

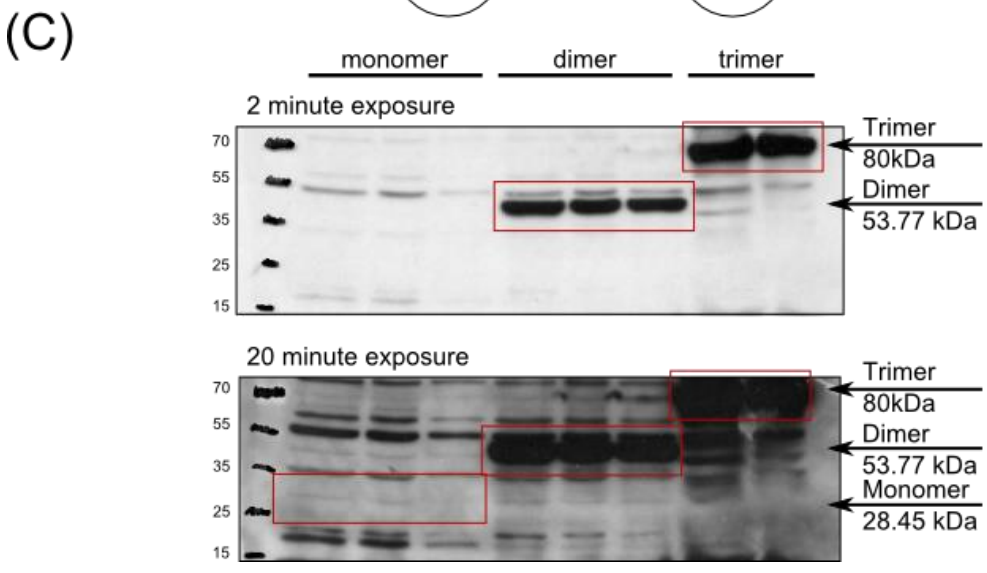
(A) Three different constructs were designed containing either one, two or three sequence optimisations of mEos2. Three different sequence optimisations (mEos2A, B and C) of mEos2 were used in order to reduce recombination. When integrated into the genome by cassette exchange, each construct will be under the control of the uracil inducible *urg1* promoter and a *spo5* DSR element.

(B) The process of creating the constructs. Three separate codon optimisations of mEos2 were designed and ordered from Genscript. The first optimisation, mEos2A, was cloned into a pAW8ENdeI plasmid, which contains the loxM and loxP sites for cassette exchange, to give the monomer construct. The second optimisation, mEos2B, was sub-cloned upstream of the first to give the dimer construct and the third optimisation, mEos2C, sub-cloned upstream of the first two to give the trimer construct. All three Plasmid constructs were amplified in *E. coli* and purified and the sequence checked. The constructs were then transformed into the *S. pombe* strain AW459 which contains the loxM-loxP exchange site at the *urg1* locus (Watson *et al.*, 2008).

(C) Two different exposures from the same Western blot of a protein extract from strains expressing the monomer, dimer and trimer constructs. Cultures were grown in EMM with supplemental uracil to induce expression. After 2 minutes exposure bands can be seen that likely correspond to the trimer (~80 kDa) and the dimer (~54 kDa), a longer exposure of 20 minutes was needed to show a band that seems to correspond to the monomer. The monomer should be ~28 kDa.


(A)
pUrg1
mEos2A
Spo5 DSR element
pUrg1
mEos2B
mEos2A
Spo5 DSR element
TGS linker
pUrg1
mEos2C
mEos2B
mEos2A
Spo5 DSR element
TGS linker
TGS linker


(B)
loxM
loxP
pAW8
loxM
monomer
loxP
loxM
loxP
loxM
dimer
loxP
loxM
loxP
loxM
trimer
loxP


(C)
monomer
dimer
trimer
2 minute exposure
70
55
35
25
15
Trimer
80kDa
Dimer
53.77 kDa
20 minute exposure
70
55
35
25
15
Trimer
80kDa
Dimer
53.77 kDa
Monomer
28.45 kDa

Imaging data for the mEos2 trimer construct was collected using pulsed activation with 300 frames at 50 ms exposure collected between each 405 nm pulse. Collecting the data in this way should mean that all blinking events for a single fluorophore happen within the interval between pulses, making it simpler to trace together repeat localisations of the same fluorophore. Spatial distribution analysis (as described in 3.2.2.1) was performed and the majority of increased density found to fall below 200 nm as in the mEos3.1 data (Figure 3.16). For this reason the data were traced using localisations within each pulse interval and distance threshold of 200 nm. The traced data were then clustered with a distance threshold of 50 nm and no time threshold and the histogram of cluster sizes analysed as described previously to predict *N* and *p* (Figure 3.16). With these thresholds the *N* value returned for trimer was 1 for a molecular density of 0.34 molecules/μm^2 (data not shown), similarly to the mEos3.1 pulsed experiment (Figure 3.14). It is possible that the difference in photo-activation efficiency between mEos3.1 and mEos2 is less in *S. pombe* than that observed by Durisic *et al.* in *Xenopus* cells. Alternatively it may be that, as suggested for the mEos3.1 trimer, cleavage or failure to properly fold means there are few or no actual trimers of, activatable, excitable mEos2. These results suggest that there is no particular benefit to using mEos2 over mEos3.1.

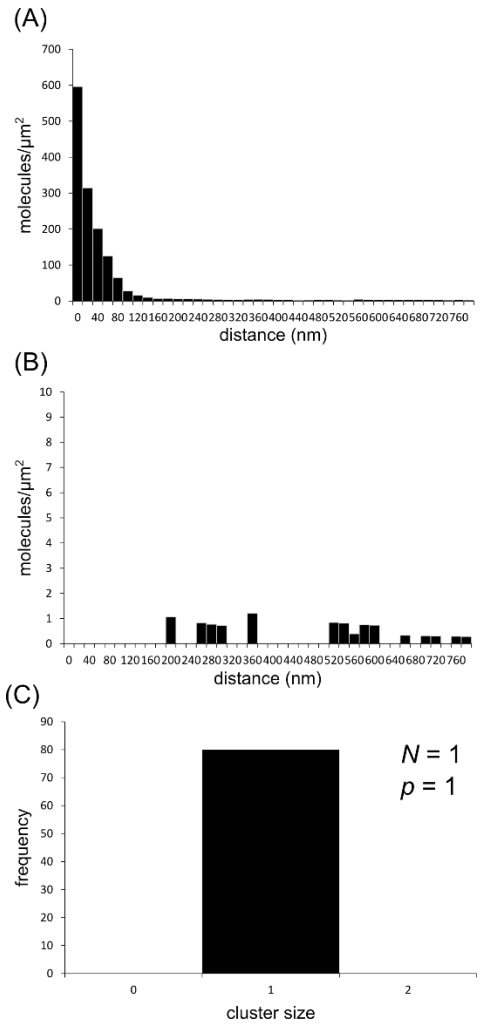


Figure 3.16 Spatial and clustering analysis of mEos2 trimer

Data were collected using pulsed activation from cells expressing a trimeric construct of mEos2.

(A) Spatial analysis of the localisation data from mEos2 shows that the majority of increased densities are below 200 nm so this is the distance to use for tracing.

(B) Since activation was pulsed and the number of frames between pulses was sufficient for the majority of activated fluorophores to be photo-bleached it should be possible to eliminate blinking by tracing within the pulse interval at the distance suggested by spatial distribution analysis. However when this was done the spatial analysis distribution histogram was flattened as would be expected for a monomer but not for trimer.

(C) Following tracing the clustering algorithm was run and the output *N* was 1. This is very similar to what was seen for mEos3.1.

3.3 Discussion

3.3.1 Implications of the simulated density limit for quantitative PALM in *S. pombe*

Simulated data have demonstrated that there is a density limit for correctly counting molecules in a cluster using PALM. This limit has implications for the expression density at which quantitative PALM is useful in *S. pombe*, and thus which proteins it can be used to study. The approximate volume of an *S. pombe* cell can be calculated by modelling it as a cylinder, *S. pombe* cells are roughly 3 μm in diameter and 15 μm long. If volume = $\pi r^2 h$ where r=radius and h=height (or length) then the approximate volume of an *S. pombe* cell is 106 μm^3 (Figure 3.17). However, what is recorded by the camera will essentially be a two dimensional projection of the focal volume. Assuming the focal plane of 500 nm is in the centre of the "cylinder" the focal volume will be roughly cuboid and therefore 15 μm × 3 μm × 0.5 μm ~22.5 μm^3 (roughly $1/6^{th}$ of the volume of the whole cell) projected onto a two dimensional area of 15 μm × 3 μm = 45 μm^2 (Figure 3.17). Results from the simulation correctly fitted the trimeric construct, given a p of 0.4 and a precision of 20 nm, at a density of 9.16 clusters/μm^2. This means that, for a trimer, the algorithm should return the correct answer given <412 clusters in the focal volume (9.16 × 45), assuming that the clusters were evenly distributed throughout the whole cell and not localised to the nucleus or other organelles. However, out-of-focus fluorescence from molecules outside of the focal plane will add to the background, reducing precision, and if an out-of-focus PSF overlaps with an in focus PSF then both localisations may be discarded during fitting. Moreover, clusters at the edge of the focal plane may have fluorophores inside the focal plane, which may be counted, and outside the focal plane, which may be discarded. These problems with projecting a three dimensional system into two dimensions for analysis mean that the density limit could be even lower in cells than the simulation suggests.

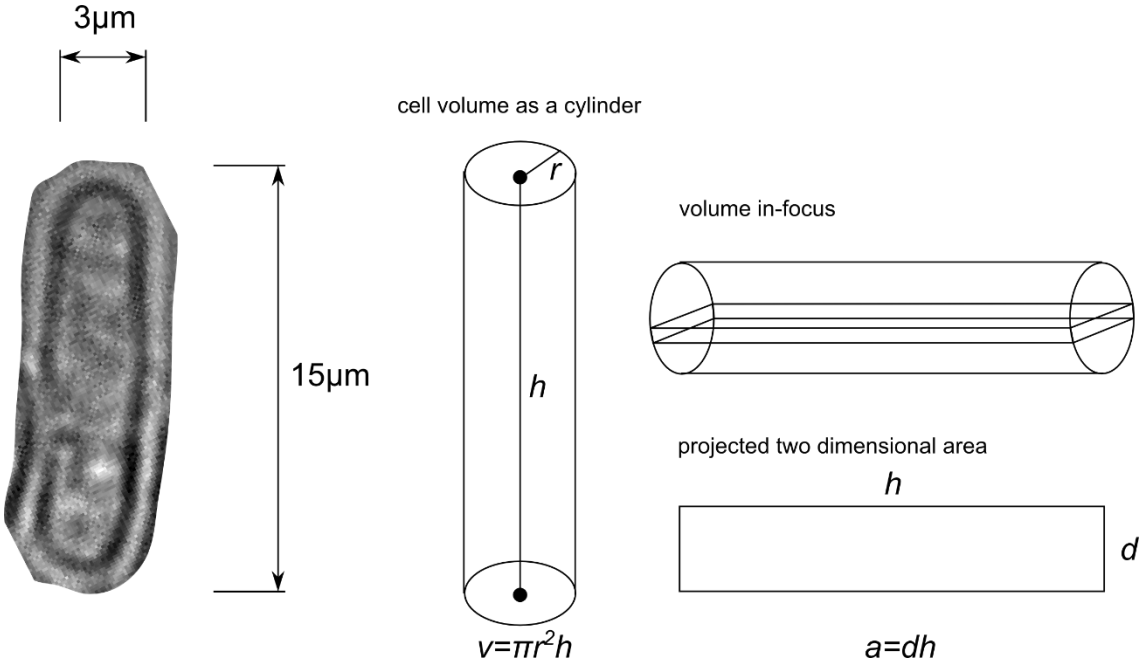


Figure 3.17 Approximating the volume of an *S. pombe* cell as a cylinder

The rod-shaped cells of *S. pombe* can be modelled as roughly cylindrical in three dimensions. However when imaging the focal depth of the microscope system is 500 nm. Assuming the focus is placed in the centre of the cell this gives a roughly cuboid volume of focus which is then projected as a two dimensional image on the camera.

Puchner *et al.*, did not cite a maximum density of clusters for their technique. Simulations suggested that counting could only be done accurately when there were fewer than 10 clusters per μm^2 for trimers of mEos3.1. Experimental data taken using a trimeric construct of mEos3.1 expressed at densities below the limit set by simulations failed to produce the expected *N* or *p* values. This failure may have been a result of failure of all three mEos3.1 units to mature or fold properly, cleavage of the trimer construct into monomers of mEos3.1 or low photo-activation efficiency of the mEos units. The results of the simulations suggest that this approach is unsuitable for anything highly expressed and densely distributed in the cytoplasm. It may be that even at sparser distributions this approach is still not usable due to the problems of long-lived dark states, background and out-of-focus fluorescence.

One of the intended applications for quantitative PALM in this study was to investigate the stoichiometry of large complexes involved in cellular processes, for example the regulatory subunit of the ribonucleotide reductase complex $Cdc22^{R1}$. It is hypothesised that $Cdc22^{R1}$ has a hexameric and a dimeric form in cells (see Introduction 1.3.2). For a hexameric cluster it would be possible to look at a density somewhat greater than for trimeric clusters due to the increased sampling (3.2.1). However, one measure of the $Cdc22^{R1}$ content of the cells suggests there could be as many as 88000 molecules per cell on average (Carpy *et al.*, 2014). Even if every molecule of $Cdc22^{R1}$ was a hexamer (i.e. excluding the dimer) this would be >14000 clusters in a single cell. The simulation results indicate it is possible to differentiate 412 clusters of 3 in the focal plane which is $1/6^{th}$ of the cell so it should be possible to differentiate a total of 2472 in a single cell, assuming they are evenly distributed. This is far below the estimated density of $Cdc22^{R1}$ suggesting this particular technique is unlikely to be able to deduce the stoichiometry of $Cdc22^{R1}$ *in vivo*.

The simulations suggest that a better photo-activation efficiency increases the molecule

density at which the clustering analysis would return the correct N. Thus there may be some benefit to investigating different PA-FPs. For example Durisic *et al*, (2014) give Dendra2 and mMaple higher photo-activation efficiencies than mEos3. One potential problem with other PA-FPs is that the mEos family is one of the brighter PA-FP families and dimmer fluorophores are likely to be fit with a worse precision (higher uncertainty). Having shown that precision is a major influence on the densities of molecules at which the algorithm can correctly determine *N* (Figure 3.7) it would be unlikely to achieve better counting results if the precision was worse.

3.3.2 Checking the sensitivity of the fitting algorithm

The results of the spatial and dark time analysis suggested that the fitting algorithm used might be too sensitive and may be picking up noise. To determine if this is the case one approach would be to try filtering with a variety of different SNRs and minimum precision settings until the dark time analysis loses the long shallow slope. It could also be useful to analyse the traces produced by the tracing algorithm to look for outlier localisations that are appearing very distant in time from anything else.

3.4 Summary

The purpose of this project was to explore the potential advantages and limitations of quantitative PALM for counting the number of subunits in protein complexes localised in the cytoplasm. Clustering simulations show that although counting single molecules in a complex in the cytoplasm is theoretically possible there are many constraints on the conditions under which it works. In particular, the precision of localisation and the density of the complexes to be studied can limit counting ability. The simulations were replicated *in*

vivo by looking at artificial oligomers of mEos3.1 expressed in the cytoplasm. It has been shown that when molecular densities that are correctly analysed in simulations are reproduced *in vivo* the system does not return the expected *N* and *p* values. Using continuous activation at low levels resulted in over-counting of cluster size (*N*) where the time threshold was short, presumably because of insufficient tracing of blinking molecules and undercounting where the time threshold was long. The *p* values did not match the expected *p* values of 0.4 for any clustering experiments but were much lower. Using pulsed activation and stringent tracing the cluster size was under-counted, possibly because of a problem with the artificial oligomers or because the fitting algorithm was fitting noise which interfered with the analysis. Repeating the analysis with mEos2, said to have higher photo-activation efficiency, did not improve the results. Taken together these results suggest that quantitative PALM is not the best technique for counting the number of subunits of cytoplasmic protein complexes.

4. Exploring FCS to determine the stoichiometry of the RNR complex

4.1 Introduction

Having determined that the density of molecules is a limitation in quantitative photo-activated localisation microscopy (PALM) it was decided that it would not be useful for answering questions about the stoichiometry of the $Cdc22^{R1}$ subunit of the ribonucleotide reductase (RNR) complex, because of the subunit's high expression level. Another single molecule fluorescence-based technique was therefore investigated. This technique is an adaptation of fluorescence correlation spectroscopy (FCS) where the number of active fluorophores can be controlled by selective activation of low numbers of photo-activatable fluorescent proteins. FCS can be used to study molecular dynamics both *in vivo* and *in vitro*. In this chapter the principles of FCS are introduced along with some of the biological applications and two means of fitting the data, the conventional method (a discrete one, two or three diffusion-component-time model) and the maximum entropy method (a probability distribution model). The varying stoichiometries of the RNR complex are also introduced in detail.

Results are presented characterising the system, using an organic dye, purified mEos3.1 and mEos3.1 expressed in cells, with the maximum entropy fitting method. Subsequently it is shown that it is possible to use FCS to differentiate between a monomer and a trimer of mEos3.1 expressed in separate cells but that it is more difficult to differentiate between a 3 kDa and a 10 kDa rhodaminated dextran in solution in a mixture. Finally, it is demonstrated that the large RNR subunit $Cdc22^{R1}$, when tagged with the mEos3.2 fluorophore, is prone to cleavage such that the tag is removed from the $Cdc22^{R1}$ protein. However, maximum entropy fitting can be used to detect a population with a long

diffusion time which may correspond to a complex and that this population cannot be seen in cells that have been treated with the RNR inhibitor hydroxyurea (HU). The results are discussed with reference to future experiments to determine whether the RNR stoichiometry varies under particular circumstances and the potential limitations of the system.

4.1.1 Fluorescence correlation spectroscopy

4.1.1.1 What is fluorescence correlation spectroscopy?

FCS is a means of measuring changes in a system via fluctuations in fluorescence emission intensity from the system over time (Lakowicz, 2006; Magde *et al.*, 1972). In FCS the excitation laser is focused as tightly as possible (ideally to a diffraction-limited spot) in the sample, giving an observable volume (sometimes called the detection volume) of ellipsoidal shape (Figure 4.1). Fluorescent molecules within the detection volume will be excited and will emit fluorescence, those outside of it will remain dark. This fluorescence signal can be captured by a detector, for example an avalanche photodiode (APD), and recorded as an intensity signal over time. A pinhole can be used to discard fluorescence from above and below the plane of focus defining the lateral cross-section of the detection volume. If the molecules within the detection volume undergo any changes that lead to a decrease in fluorescence output, or if molecules move into or out of the detection volume, this will be recorded as a fluctuation by the detector. From these fluorescence fluctuations information such as reaction kinetics and diffusion rates can be calculated. In combination with labelling with organic dye or tagging with fluorescent proteins this information can be used to study single molecules in biological systems.

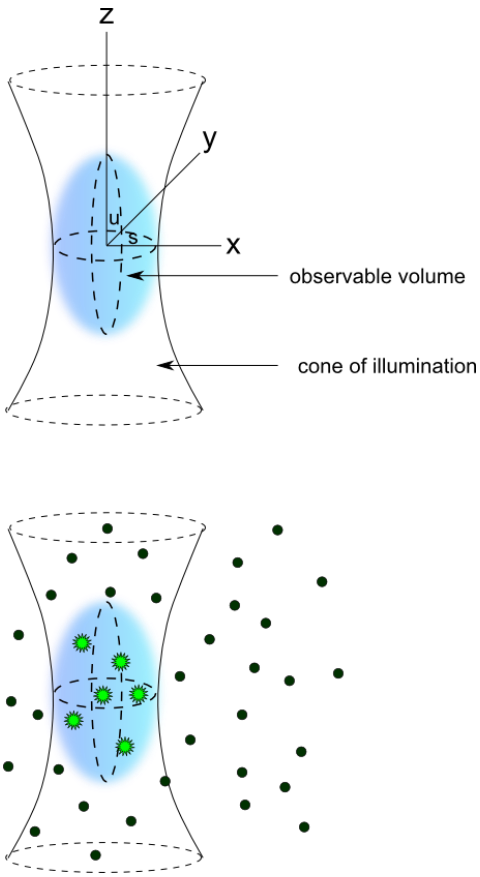


Figure 4.1 The FCS cone of illumination

Adapted from Lakowicz, (2006). As the laser is focused down to a diffraction-limited point it creates a cone of illumination with a defined volume in which fluorophores will be excited and therefore observable. This volume has defined proportions in the *x*, *y* and *z* planes and is ellipsoidal. The size of the volume will depend on how well the laser beam has been focused. The ellipsoidal volume shown does not have a hard edge; the radius *s* and the half height *u* are the distances at which the illumination profile decreases to e^{-2} of its maximum value at I_0. The equation *4-1* describes the three dimensional Gaussian distribution of the detection profile.

$$p(r) = I_0 exp[-2(x^2 + y^2)/s^2]exp(-2z^2/u^2) \qquad 4\text{-}1$$

As such some molecules may fluoresce outside of what is technically in-focus for the detector.

The technique of FCS relies on the number of excited molecules within the volume being such that each fluorophore has contributed substantially to the fluorescence signal. If the fluorophores are diffusing freely within the sample then the number of molecules within the detection volume will vary over time. The number of fluorophores within the detection volume can be described by the Poisson distribution (equation *4-2*) for which *P(n,N)* is the probability of *n* fluorophores being present in the detection volume given the average number of molecules *N*:

$$P(n,N) = \frac{N^n}{n!}e^{-N} \qquad 4\text{-}2$$

Changes in *n* will lead to fluctuations in the detected intensity. If the molecules are diffusing rapidly the intensity will change rapidly and likewise if the molecules are diffusing more slowly the intensity changes will be slower (Figure 4.2). Diffusion coefficients can be calculated based on the fluctuations in intensity being detected. Diffusion speed can also be related back to molecular size for information on binding or changes in oligomeric state.

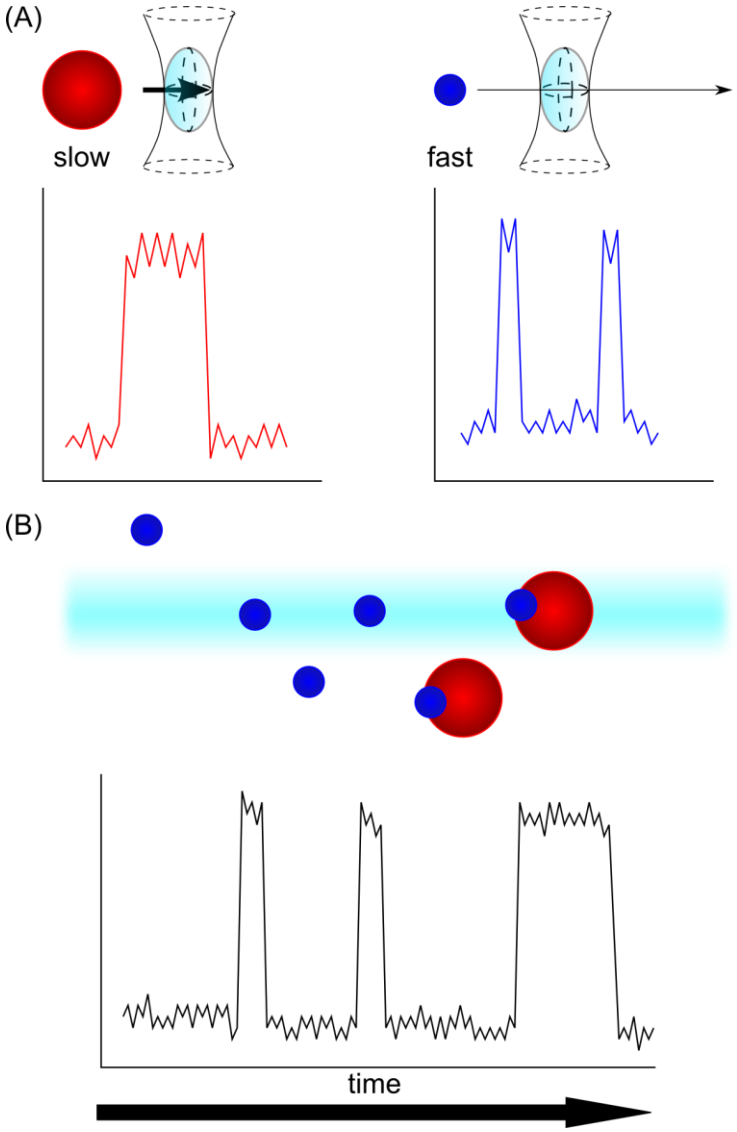


Figure 4.2 Fluctuations in intensity relate to diffusion coefficient and size

(A) When a molecule is moving slowly the intensity recorded by the detector will change slowly, when a molecule is moving fast the change in intensity will be fast. If the molecules observed are diffusing freely then there will be a relationship between diffusion speed and the size of the molecule.

(B) If there is a change in the size of the molecule being observed, for example a small fluorescent ligand binding to a large protein, then a change will be observed in the way the fluorescence intensity is fluctuating.

4.1.1.2 The invention of FCS

FCS was first described in a series of papers in the early 1970s. Magde *et al.*, (1972) used it to demonstrate the binding kinetics of EtBr with DNA by measuring fluorescence fluctuation in an excited volume and analysing it using an autocorrelation function. EtBr is only fluorescent when in complex with DNA so Magde *et al.*, reasoned that fluctuations in the fluorescence intensity could be directly related to fluctuations in the number of DNA-EtBr complexes. The same group published a further paper describing how FCS can be used to measure kinetic constants and diffusion coefficients (Elson and Magde, 1974; Magde *et al.*, 1974). They demonstrated this technique by looking at pure diffusion of a fluorescent molecule in the absence of any other chemical reaction, the binding of a small, rapidly diffusing ligand to a larger binding partner, and a uni-molecular isomerisation. A year later a study described the way Brownian motion and rotational diffusion (two different types of particle movement in solution) relate to the correlation function in FCS (Aragón and Pecora, 1975). However all of the earliest studies suffered from a poor signal-to-noise ratio (SNR) as a result of low detection efficiency, large numbers of molecules within the excited volume, unstable laser sources and fluorophores with a low quantum yield. FCS measurements are best performed on molecules at a low concentration such that few are being excited at the same time, typically in the 1 nM range. Since the samples are so dilute it is necessary to supress background fluorescence and scattering by using confocal optics and appropriate filtering. These requirements meant that FCS was not widely used for some time after its conception.

Once confocal optics started to become more widely and economically available FCS started to become a more practical technology (Rigler *et al.*, 1993). In addition high efficiency APD detectors and more stable lasers improved the quality of SNR achievable. This opened up the field to a greater number of people and allowed a greater exploration

of what FCS could be used for, eventually leading to studies using FCS in living cells (Berland *et al.*, 1995; Politz *et al.*, 1998; Rigler *et al.*, 1999).

4.1.1.3 The biological applications of FCS *in vivo*

Berland *et al.*, (1995) were the first to utilise FCS in the cytoplasm of a cell, exploiting the confocal optics that had recently become available. They initially used FCS to measure the diffusion coefficients of various particles of known size in solutions of varying viscosities. They then went on to show that they could use it to calculate the diffusion coefficients of two different radius latex beads (7 nm and 15 nm) that had been injected into mouse fibroblasts. They demonstrated that they obtained a different diffusion coefficient for each radius of bead and that the intracellular diffusion rates were similar to, but slightly slower than, free diffusion in water. Similarly Politz *et al.*, (1998) were able to utilise FCS in the nucleus of rat fibroblasts injected with fluorescein labelled oligo deoxy-nucleotides to show that the movement of the oligos within the nucleus was comparable to their movement in water.

One of the ways in which FCS can be used to study biological phenomena is that the information it gives on diffusion rates can reveal something about the sizes of the particles diffusing. In the case of small peptide ligands binding to large receptors this can give information about reaction rates or the size of complex being formed. The earliest *in vivo* work on studying binding by FCS was largely performed in the membrane, for example Rigler *et al.*, (1999) were able to study the binding properties of Rhodamine labelled proinsulin C-peptide to the membranes of various human cell lines *in vivo*. C-peptide having recently been shown to have beneficial effects in animals with experimental diabetes and human patients with diabetes type 1. At a similar time, Korlach *et al.*, (1999) were using FCS to measure changes in the diffusion coefficient of DiI-C_{20} (a fluorescent probe with

an acyl tail that can diffuse freely in a membrane) in lipid phases that were fluid, highly ordered or containing a high level of cholesterol. The FCS was able to distinguish two different diffusion coefficients in a region where two different phases coexisted on a single vesicle, which had not been distinguishable using fluorescence microscopy. Zhong *et al.*, (2001) used Rhodamine-labelled insulin to study insulin binding to human renal tubular cells. They found two different diffusion coefficients suggesting two classes of binding site which agreed with a previous experiment done using a radio-ligand binding assay, but they found that the FCS gave additional data that allowed a more precise evaluation. Patel *et al.*, (2002) were even able to use FCS to study how ligand binding causes heptahelical receptors to oligomerise in the membrane of CHO-K1 cells. Subsequently researchers have utilised FCS for a variety of biological purposes (reviewed in Chen *et al.*, 2008; Kim and Schwille, 2003).

4.1.1.4 Calculating the autocorrelation function

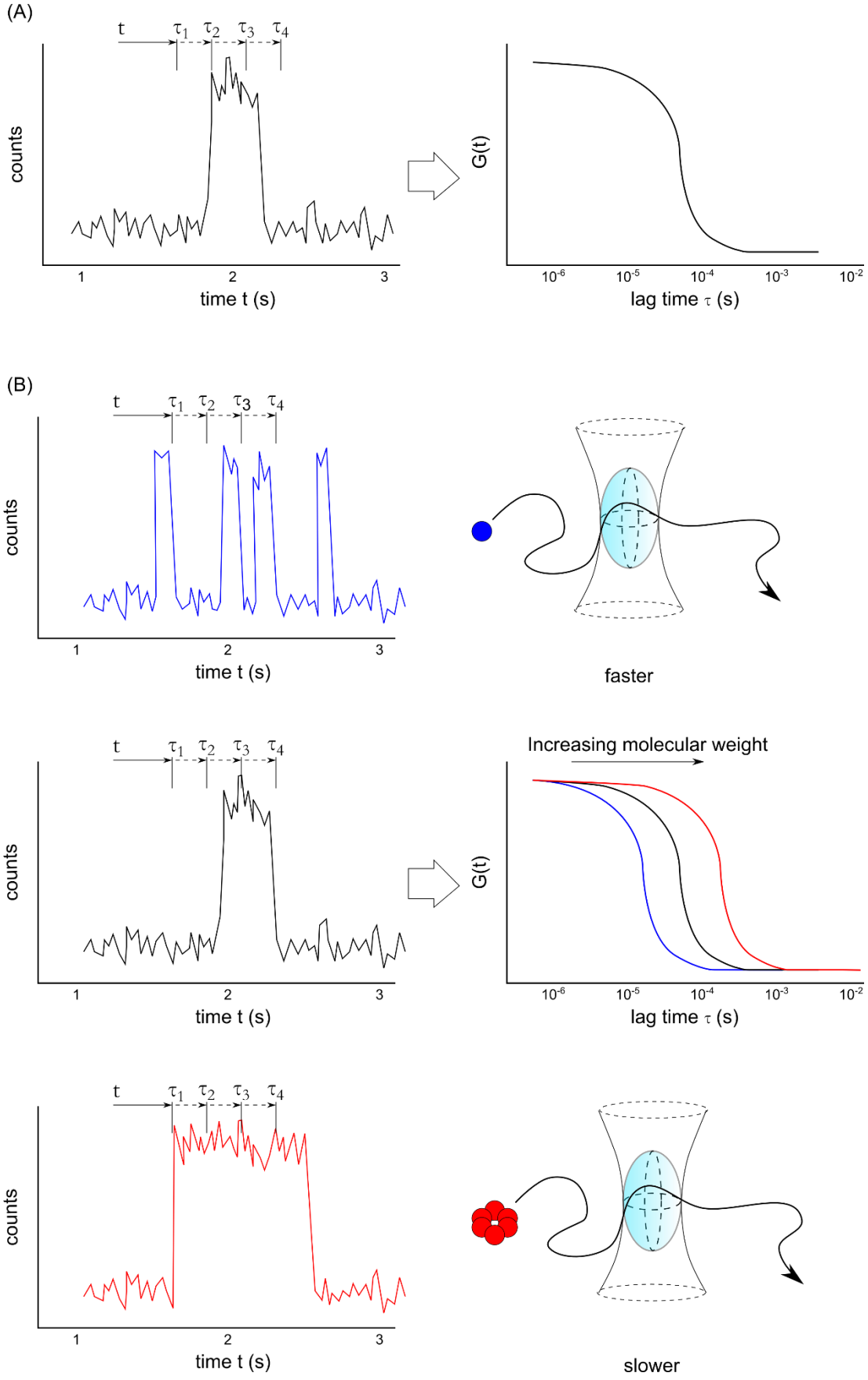
To analyse the intensity fluctuations autocorrelation is performed on the raw intensity data. Autocorrelation is the correlation of a value from a series with other values from the same series separated by a time interval (the lag-time τ). In FCS the intensity at a given time $F(t)$ is compared to the intensity at a later time $F(t+\tau)$. The correlation amplitudes $F(t)F(t+\tau)$ are calculated and summed for all t over a range of lag-times τ to give the autocorrelation function $G(\tau)$. $G(\tau)$ can then be plotted against τ to give a decay curve, hence referred to as the autocorrelation curve. In general the correlation will be high for small values of τ and will decrease as τ increases. If diffusion of the observed molecules is fast the correlation will start to decrease at smaller τ values than if diffusion is slow (Figure 4.3).

Figure 4.3 Relationship between intensity fluctuation and autocorrelation curve

In FCS the detector records fluorescence emission from the detection volume as a photon count which can be output as a trace of counts over time. Any change in the fluorescence from molecules within the detection volume will cause a fluctuation in the counts. For example, if the fluorescent molecules are diffusing freely the number of molecules within the volume will change, affecting the recorded counts. This count fluctuation can then be run through the autocorrelation process to give a decay curve. Various information can be extracted from this autocorrelation curve, for example the diffusion coefficient.

(A) The autocorrelation function (*G(t)*) compares the intensity recorded at two times points separated by the time value τ (t and $t+\tau$) for a variety of values of τ. When *G(t)* is plotted against lag time (τ) it gives a decay curve.

(B) If the molecule is moving fast then the intensity trace will be fluctuating rapidly and the correlation between point t and point $t+\tau$ will decay more sharply. Whereas if the molecule is moving slowly the intensity trace will fluctuate slowly and the correlation between point t and point $t+\tau$ will decay more slowly. In the autocorrelation curve this will appear as a rightward shift for decreasing speed.


(A)
t
τ1 τ2 τ3 τ4
counts
1
2
3
time t (s)
G(t)
10-6
10-5
10-4
10-3
10-2
lag time τ (s)
(B)
t
τ1 τ2 τ3 τ4
counts
1
2
3
time t (s)
faster
t
τ1 τ2 τ3 τ4
counts
1
2
3
time t (s)
Increasing molecular weight
G(t)
10-6
10-5
10-4
10-3
10-2
lag time τ (s)
t
τ1 τ2 τ3 τ4
counts
1
2
3
time t (s)
slower

The autocorrelation function contains information on the diffusion coefficient and the number of molecules within the detection volume. The value of the autocorrelation function for a time delay of τ is given by the average value of the product amplitudes $F(t)F(t+\tau)$ (equation *4-3*) where T is the data collection time. The value t refers to real time and τ refers to the lag-time difference between two intensity measurements.

$$G(\tau) = \langle F(t)F(t+\tau)\rangle = \frac{1}{T}\int_0^T F(t)F(t+\tau)dt \qquad 4\text{-}3$$

Where $\langle F \rangle$ is the average intensity, the fluctuations of $F(t)$ around the mean value can be described as:

$$\boldsymbol{\delta F(t) = \langle F\rangle - F(t)} \qquad 4\text{-}4$$

Such that the autocorrelation function for the fluorescence intensities, as normalised by the square of the average intensity can be given by equation *4-5* (Lakowicz, 2006).

$$G(\tau) = \frac{\langle \delta F(0)\delta F(\tau)\rangle}{\langle F\rangle^2} \qquad 4\text{-}5$$

The autocorrelation function can be plotted against τ and this curve can be fitted with a model of the system being observed to enable calculation of the average residence time in the detection volume.

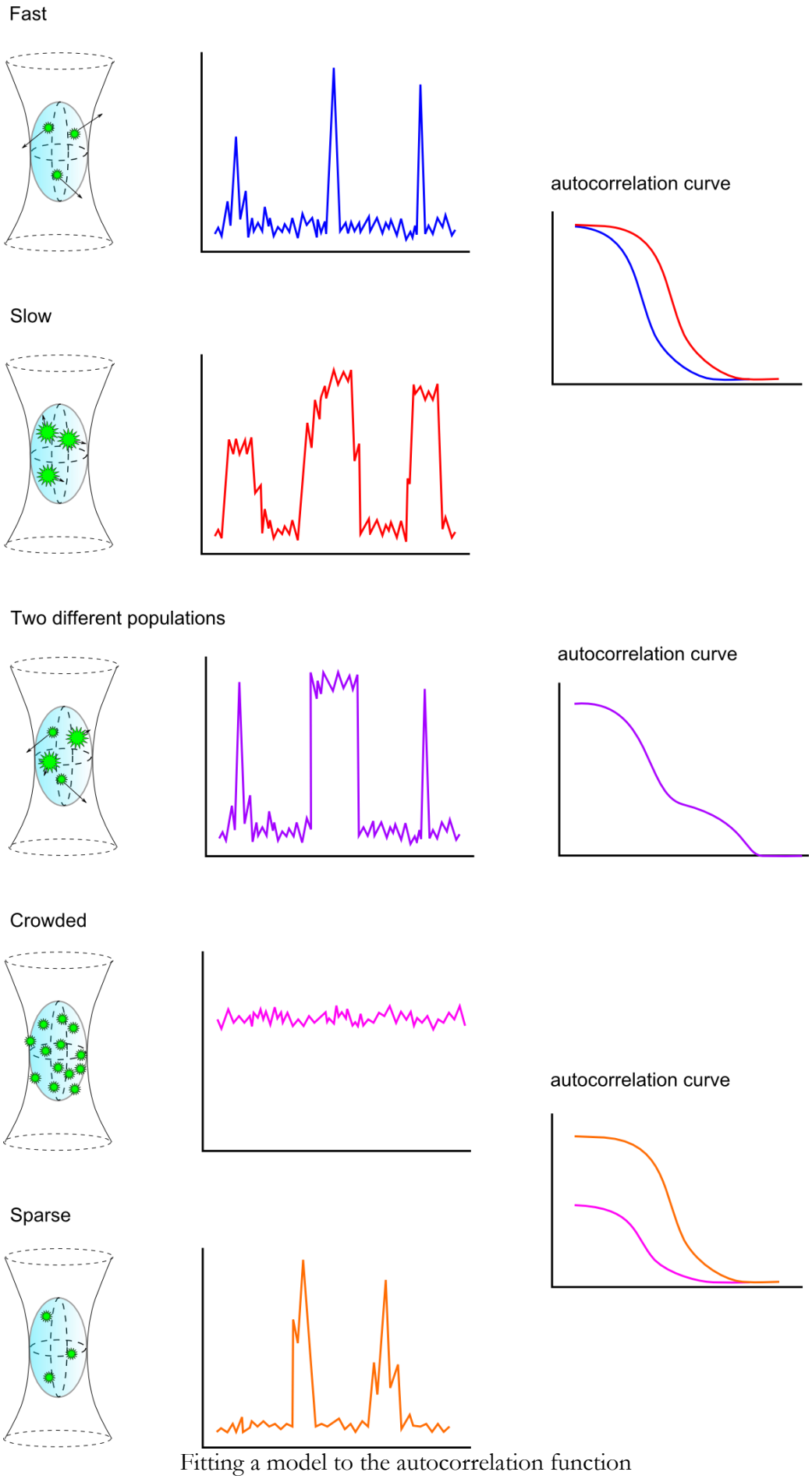
The autocorrelation curve for a single species should show a rightward shift for increasing molecular weight (or strictly speaking increasing hydrodynamic radius of the molecule) (Figure 4.4). As diffusion speed varies with the cubed root of the molecular mass the generally accepted rule is that a 3-5 fold increase in molecular weight would be required to see a difference in residence time that can be distinguished by FCS (Haustein and Schwille, 2007; Meseth *et al.*, 1999). Additionally the correlation amplitude $G(0)$ at lag-time zero is inversely proportional to the value N, the average number of molecules in the detection volume. Thus the smaller N is the larger the correlation amplitude and, in general, the

clearer the autocorrelation curve. The caveat to this is that longer acquisitions or more accumulations must be done in order to have sampled sufficient molecules. The autocorrelation function of a sample with multiple populations is the sum of the autocorrelation decay curves for each individual population and as such the presence of two differently diffusing populations will change the shape of the autocorrelation decay curve (Figure 4.4).

To extract information on the system properties it is necessary to fit the autocorrelation decay curve with a model. For example, if the protein being studied can bind to other molecules to form a complex, two or more species will exist with different diffusion coefficients (proportional to the difference in hydrodynamic radius between the protein and the complex). This could be a fluorescently labelled ligand which binds to a larger protein; the ligand alone should diffuse faster than the ligand-protein complex, assuming no interactions with static molecules such as the membrane. The model would need terms to represent both these components. The model used to fit the autocorrelation curve varies depending on how many components the system is assumed to have, whether they are modelled as diffusing in two or three dimensions and whether they are found in equal or different proportions to each other.

Figure 4.4 Information contained in the autocorrelation curve

The autocorrelation curve will vary depending on the properties of the molecule(s) in the system being studied. In particular it contains information on the diffusion speed of the molecule being observed and the number of those molecules in the system. If two molecules of different speeds are compared by FCS, a shift occurs in the position of the autocorrelation curve. For a system containing more than one component the curve will be the sum of the two individual curves for each component. The correlation amplitude at *G(0)* of the autocorrelation curve is inversely proportional to the average number of fluorescent molecules in the detection volume, meaning the less dense the molecules in the system are, the higher the correlation amplitude will be.



Fitting a model to the autocorrelation function

4.1.1.4.1 Fitting with a conventional model

The conventional way of fitting an FCS autocorrelation decay curve is with an equation that assumes a system containing a number of species of discrete diffusion speeds. A model is used which also assumes the number of components in the system. This method is limited to around 3-4 different species in a system (Lakowicz, 2006).

The autocorrelation function for a system containing a single component, diffusing in three dimensions, in a homogenous medium is given by equation *4-6* where *N* is the average number of molecules in the detection volume and k^2 is the ratio of the axial to radial dimensions of the mean square intensity of the focal volume (Chen *et al.*, 2008). The value τ_D is the diffusion time, which is the transit time of the molecule across the detection volume.

$$\boldsymbol{g(t) = \frac{1}{N_p}\frac{1}{(1+\tau/\tau_D)}\frac{1}{(1+\tau/k^2\tau_D)^{1/2}}} \qquad 4\text{-}6$$

Where *G(t)* from the previous equations is the experimental autocorrelation function and *g(t)* represents an analytical functional form for fitting. For a two component system with two populations at unequal concentrations with differing diffusion speeds the autocorrelation function is a product of the two separate autocorrelation functions.

$$g(t) = \frac{g_1(t)\langle N_1\rangle^2 + g_2(t)\langle N_2\rangle^2}{(\langle N_1\rangle + \langle N_2\rangle)^2} \qquad 4\text{-}7$$

N_1 is the number of diffusing particles of component 1 and N_2 is the number of diffusing particles of component 2. If the two g_1 and g_2 values are substituted for each component's autocorrelation function the equation becomes:

$$g(t) = \frac{1}{N_p}\left\{\frac{f_1}{\left(1 + {}^{t}/_{\tau_{D1}}\right)\left(1 + {}^{t}/_{k^2\tau_{D1}}\right)^{1/2}} + \frac{f_2}{\left(1 + {}^{t}/_{\tau_{D2}}\right)\left(1 + {}^{t}/_{k^2\tau_{D2}}\right)^{1/2}}\right\} \qquad 4\text{-}8$$

Where f_1 and f_2 are the respective fractions of each component (sometimes represented as f and $1\text{-}f$). It is possible to model a system of more than two components in this way. For a multi-component system the equation becomes:

$$g(t) = \frac{1}{N_p}\sum_{i=1}^{M}\frac{f_i}{(1 + t/\tau_{Di})(1 + t/\kappa^2\tau_{Di})^{\frac{1}{2}}} \qquad 4\text{-}9$$

The diffusion coefficient can be extracted from the diffusion time value τ_D given the radius of the excitation beam.

$$D = \frac{\omega^2}{4\tau_D} \qquad 4\text{-}10$$

Where ω is the radius of the detection volume, D is the diffusion coefficient and τ_D is the diffusion time (Chen *et al.*, 2008; Lakowicz, 2006).

4.1.1.4.2 Fitting with a maximum entropy distribution

Whilst the standard model is very useful for calculating diffusion coefficients from the autocorrelation function, it is very difficult to apply it to systems with more than three distinct populations. It is also less useful when looking at populations with very similar diffusion coefficients. The maximum entropy distribution fitting method was first suggested for analysis of FCS data by Sengupta *et al.*, (2003), who noted that the standard method of fitting FCS data was only adequate for systems with a small number of discrete populations. The benefit of the maximum entropy fitting method is that it makes no assumptions about the number of populations within the system or whether their diffusion

speeds are discretely distributed. The principle of maximum entropy is that the probability distribution which best represents a system is the one which has the maximum entropy, within the constraints of prior data (Skilling and Bryan, 1984). Sengupta *et al.*, were able to show that their maximum entropy fitting algorithm produced similar results to the conventional fitting in a simple, single component system and fitted the data better in the case of a large number of similar components (Figure 4.5). With the maximum entropy distribution the x axis is the autocorrelation time value τ and the area under the graph is the probability that a component falls in that distribution.

As it is possible that $Cdc22^{R1}$ exists in more than one oligomeric state (see 4.1.2.1), the system of interest may be somewhat heterogeneous and it is not possible to know how many components there may be. Therefore, the maximum entropy fitting may be useful for analysing the data.

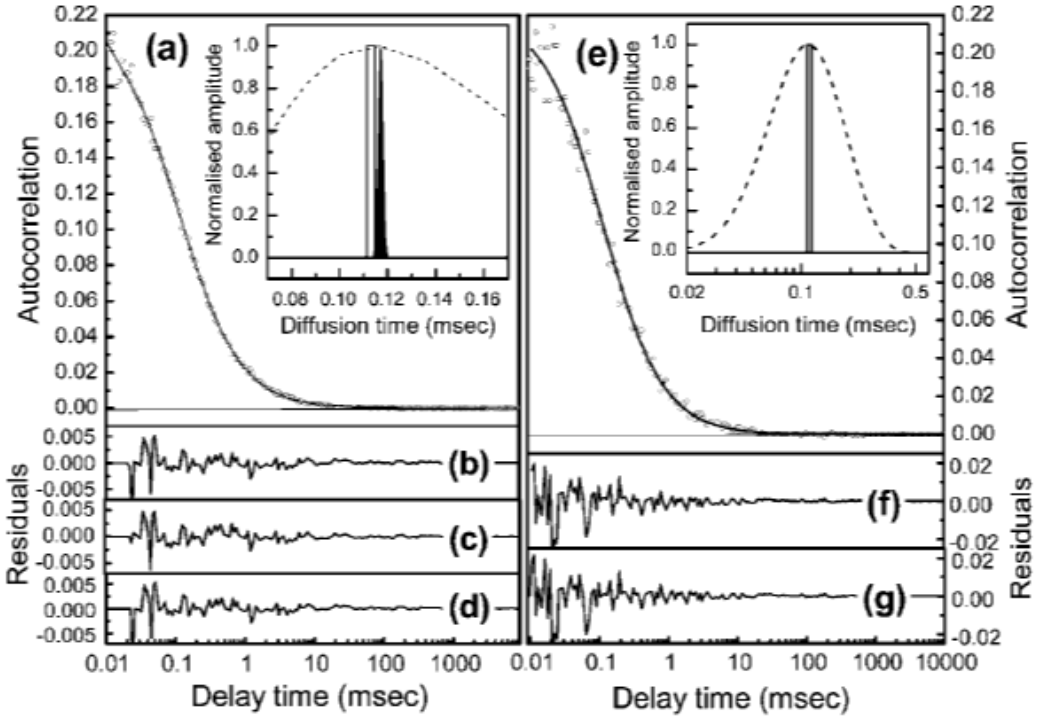
Figure 4.5 Comparison of the maximum entropy fitting algorithm with conventional fitting

Figures taken from Sengupta *et al.*, (2003).

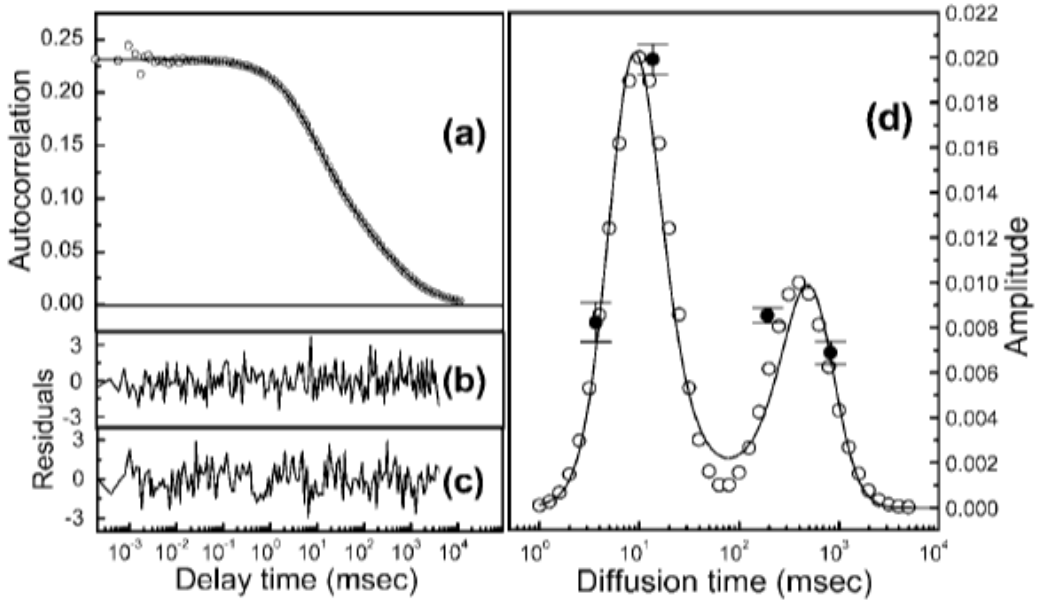
(A) Sengupta *et al.*, demonstrated that in a simple single component system of EGFP diffusing in a buffer, their maximum entropy fitting algorithm provides a similar answer to conventional fitting. The main part of (a) shows the autocorrelation data fit with both the maximum entropy algorithm and a conventional single component fit. The inset shows the distributions of τ_D as calculated with a linear fit (solid line), a maximum entropy fit with a logarithmic distribution of τ_D (black dashed line), and the maximum entropy fit with a narrow linear distribution of τ_D (shaded area). The same is shown in (e) but for data with a lower signal to noise ratio.

(B) They then demonstrated that for simulated data of a more heterogeneous system the maximum entropy fit gave a better description of the system. The simulated data contained 41 components, residuals are for (b) the maximum entropy fit and (c) the conventional fit. In (d) the input simulation data are plotted as open circles and the maximum entropy distribution as a solid line, four component conventional fit data are plotted as filled circles with error bars. It can be seen that the maximum entropy distribution follows the input data quite closely, whereas the four component fit clusters its components around the peaks.

(A)


(a)
(b)
(c)
(d)
(e)
(f)
(g)
Autocorrelation
Residuals
Normalised amplitude
Diffusion time (msec)
Delay time (msec)

(B)


(a)
(b)
(c)
(d)
Autocorrelation
Residuals
Amplitude
Delay time (msec)
Diffusion time (msec)

4.1.1.5 Detectors used in FCS

An APD is a semiconductor electronic device that can be used to detect light and convert it into an electronic impulse. The avalanche effect allows the signal to be amplified within the detector. As a result APDs can have single molecule sensitivity. In the FCS system used for this study a single photon avalanche photodiode (SPAD) was used as the detector.

The difference between SPADs and other APDs is that SPADs are able to operate with a reverse-bias voltage above the breakdown voltage ("Single-photon avalanche diode," 2014). The upshot of which is that a single photon can trigger a self-sustaining avalanche cascade of photoelectrons to create a signal without front-end amplifiers. The SNR of a SPAD is limited only by shot noise, which is an inherent property of light, and it has no readout noise or dark current noise, as described in equation *4-11* where N_{shot} is the shot noise, G is the electron multiplication gain, η is the quantum efficiency, τ is the integration time and ϕ is the photon flux.

$$SNR_{SPAD} = \frac{G\eta\phi\tau}{\sqrt{(N_{shot})^2}} \qquad 4\text{-}11$$

Additionally SPADs can provide a time resolution of tens of picoseconds, much faster than an electron multiplied charge coupled device camera, the type of detector used in high resolution fluorescence microscopy (such as PALM).

4.1.2 The ribonucleotide reductase complex

The RNR complex catalyses the reaction converting ribonucleotide diphosphates (NDPs) to deoxy-ribonucleotide diphosphates (dNDPs) which are the precursors to deoxy-ribonucleotide triphosphates (dNTPs). Imbalances in the available pools of these dNTPs can lead to mutations and other genotoxic events (reviewed in Mathews, 2006), which is why dNTP levels are tightly regulated. Class Ia RNR complexes, such as the one used by *S.*

pombe, are typically made up of two subunits designated R1 (the large regulatory subunit) and R2 (the small catalytic subunit) (Nordlund and Reichard, 2006). The R1 and R2 subunits are also referred to as α and β when discussing complex stoichiometry, for example a hetero-tetrameric complex consisting of two R1s and two R2s would be designated $\alpha_2\beta_2$. In *S. pombe* the R1 and R2 subunits are called $Cdc22^{R1}$ and $Suc22^{R2}$ respectively (Sarabia *et al.*, 1993).

The R1 subunit has two allosteric binding sites for nucleotide effectors, which are involved in the regulation of NDP reduction. One site is termed the "overall activity" site (sometimes the a-site) and the other is termed the "specificity" site (sometimes the s-site) (Eriksson *et al.*, 1979). The activity site can bind either ATP, which stimulates RNR activity, or dATP, which inhibits RNR activity (reviewed by Nordlund and Reichard, 2006). The specificity site can bind ATP, dATP, dTTP or dGTP and the nucleotide bound will affect which substrate the RNR can reduce (reviewed by Nordlund and Reichard, 2006) (see Introduction 1.3.2 for more detail). It has been shown that the transcript level of the R1 subunit is regulated in a cell cycle dependent manner, with the highest transcript level occurring around G1/S phase (Björklund *et al.*, 1990; Gordon and Fantes, 1986). However this does not seem to be the only way in which RNR activity is regulated.

4.1.2.1 Stoichiometric regulation of the ribonucleotide reductase

Within the class Ia group there is evidence to suggest that the RNR complex exists in more than one stoichiometry. These stoichiometries may have different levels of activity, possibly with some forms of the complex being completely inactive. The class Ia RNR complex stoichiometry has been studied in multiple organisms including mice, humans, fission yeast and *E. coli* and a variety of techniques have been employed.

Using mouse cell extract Kashlan *et al.*, (2002) used sedimentation velocity and varying

concentrations of ATP (the RNR activator). They showed that the R1 protein was found at sedimentation times suggesting sizes that corresponded to a monomer, a dimer, a tetramer and a hexamer. Increasing levels of ATP pushed the reaction towards the higher order complexes. This demonstrated that as well as regulating activity the nucleotide effectors regulate the stoichiometry of the RNR complex and suggested a possible link between stoichiometry and activity. As ATP is an activator of RNR it was speculated that one of the higher order complexes would be the active form.

Looking at mouse cell proteins, using a measure of particle size called gas phase electrophoretic-mobility macromolecule analysis (GEMMA), Rofougaran *et al.*, (2006) also saw a monomer of R1 in the absence of effector nucleotides. At high concentrations of both dATP (inhibitor) and ATP (activator) they saw a large complex of ~510 kDa which they hypothesised to be a hexamer of R1 subunits. They were further able to show that this 510 kDa complex interacts with a dimer of the R2 subunit to give an $\alpha_6\beta_2$ complex. In solution they found that, as ATP concentration increased, there was an increase in the concentration of the 510 kDa putative R1 hexamer complex and an increase in enzyme activity. They concluded that the $\alpha_6\beta_2$ form is the major, active form of mammalian RNR. Having found that dATP also promotes R1 hexamer formation they hypothesised that an inactive R1 hexamer might also exist, becoming active when dATP is depleted and ATP is bound instead (Figure 4.6 A).

The same group then went on to look at the *E. coli* class Ia RNR proteins. Using GEMMA the *E. coli* R1 was found to be predominantly monomeric in the absence of effectors and to form a dimeric complex with 50 mM of either dATP (activator: can bind either effector site), dTTP or dGTP (can only bind the specificity site) (Rofougaran *et al.*, 2008). Using both the *E. coli* R1 and R2 they observed a complex with a size of 510 kDa that formed in the presence of dATP or dTTP + ATP but not in the presence of dTTP alone. Since

dTTP cannot bind the activity site this suggested that the activity site is involved in the formation of higher order complexes. By repeating the experiment with R2 fused to another (unrelated) protein, in order to change the size, they were able to suggest that the 510 kDa particle they saw was an $\alpha_4\beta_4$ complex and nothing larger was seen. Using enzyme activity assays they were able to draw the conclusion that the $\alpha_4\beta_4$ complex is an inactive form of RNR in *E. coli* (Figure 4.6 B).

More recently Fairman *et al.*, (2011) examined oligomerisation of the human R1 (hR1) protein using gel filtration chromatography and dynamic light scattering (Fairman *et al.*, 2011). Like Rofougaran *et al.*, they found that in the absence of effectors hR1 was a monomer, at 5 μM dATP a mixture of dimers and hexamers were formed and at 20 μM dATP the hexameric form was favoured. They also showed the same was true for *S. cerevisiae*. They did not observe anything to suggest a tetrameric hR1 complex. Using an RNR activity assay they showed that, at concentrations of dATP that should lead to mainly hexameric hR1 RNR, activity was reduced. From this they concluded that there exists an inactive hexameric form. They also used electron microscopy on the purified RNR complexes and what they saw seemed to support the theory of an R1 hexamer (Figure 4.6 C).

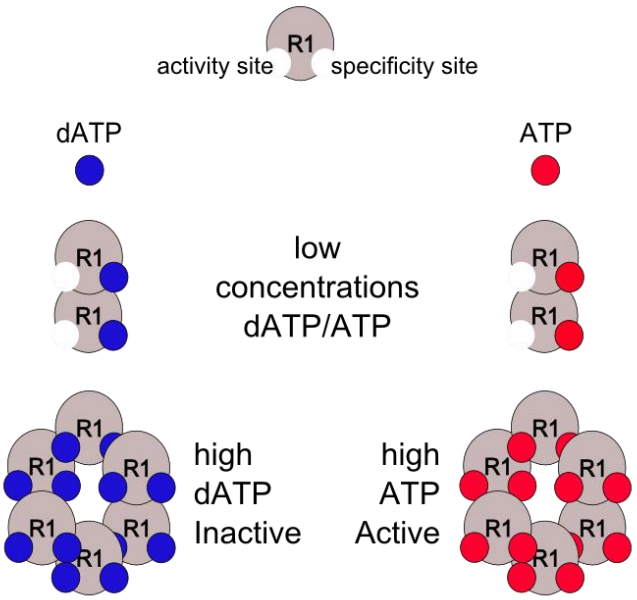
Figure 4.6 Stoichiometry in regulation of the RNR complex activity

(A) The possible stoichiometries of mouse the RNR R1 subunit according to Rofougaran *et al.*, (2006). The mouse R1 subunit of the RNR complex has one substrate binding site and two allosteric effector binding sites. ATP and dATP can bind to both effector binding sites. In the absence of these effectors the mouse R1 protein exists as a monomer. With only the specificity site occupied a dimeric R1 is induced to form (Ingemarson and Thelander, 1996). If both effector binding sites are occupied the R1 protein forms a hexamer (Rofougaran *et al.*, 2006).

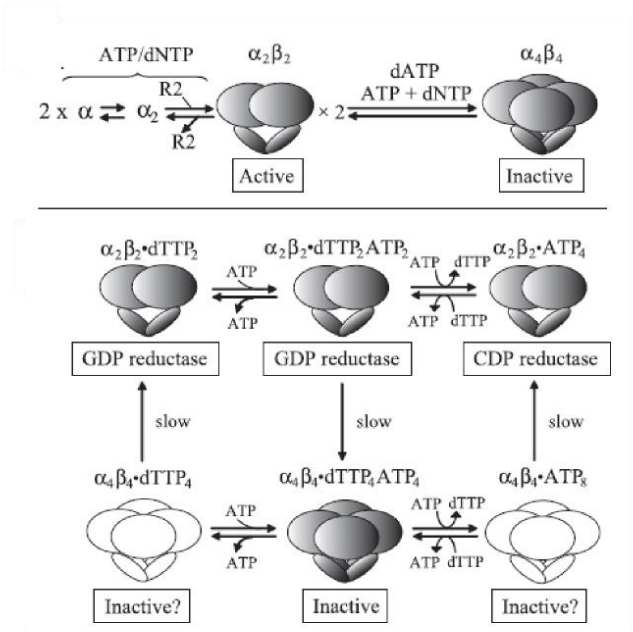
(B) Taken from Rofougaran *et al.*, (2008) a hypothetical model for the regulation of *E. coli* RNR stoichiometry. The *E. coli* RNR complex has been shown to have an active $\alpha_2\beta_2$ form and an inactive $\alpha_4\beta_4$ form, the formation of which are regulated by the binding of dATP/ATP/dTTP (Rofougaran *et al.*, 2008).

(C) Fairman *et al.*, (2011) combined electron microscopy and X-ray crystallography of the *S. cerevisiae* R1, bound to various effector molecules, to come up with a model of how RNR stoichiometry changes with different effectors. They saw evidence for monomeric, dimeric and hexameric R1. The ratio of the larger complexes increased as the concentration of dATP increased.

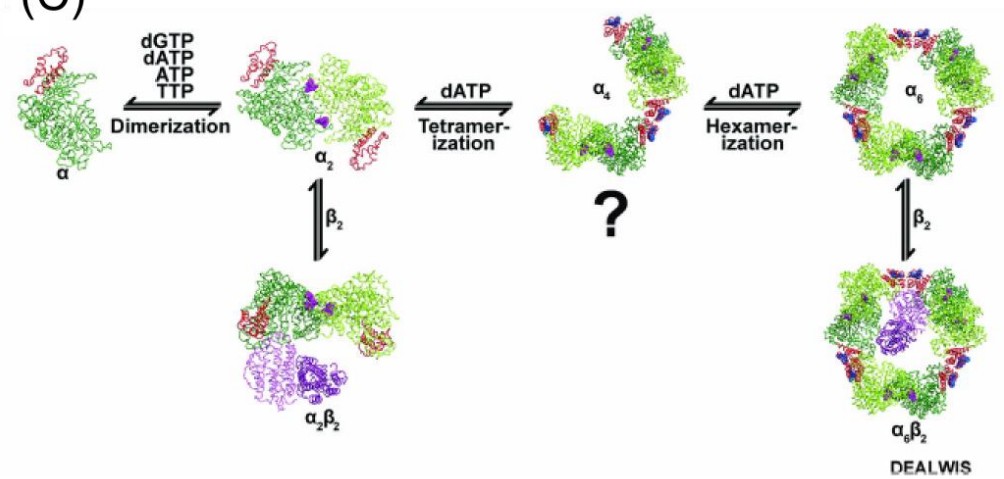
(A)


R1
activity site
specificity site
dATP
ATP
R1
R1
low
concentrations
dATP/ATP
R1
R1
R1
R1
R1
R1
R1
high
dATP
Inactive
high
ATP
Active

(B)


ATP/dNTP
$\alpha_2\beta_2$
$\alpha_4\beta_4$
dATP
ATP + dNTP
2 x α
α_2
R2
R2
× 2
Active
Inactive
$\alpha_2\beta_2$•dTTP$_2$
$\alpha_2\beta_2$•dTTP$_2$ATP$_2$
$\alpha_2\beta_2$•ATP$_4$
ATP
ATP
ATP
dTTP
ATP
dTTP
GDP reductase
GDP reductase
CDP reductase
slow
slow
slow
$\alpha_4\beta_4$•dTTP$_4$
$\alpha_4\beta_4$•dTTP$_4$ATP$_4$
$\alpha_4\beta_4$•ATP$_8$
ATP
ATP
ATP
dTTP
ATP
dTTP
Inactive?
Inactive
Inactive?

(C)


dGTP
dATP
ATP
TTP
Dimerization
α
α_2
dATP
Tetramer-
ization
α_4
?
dATP
Hexamer-
ization
α_6
β_2
β_2
$\alpha_2\beta_2$
$\alpha_6\beta_2$
DEALWIS

The model organism the Carr research group works with is the yeast *S. pombe* (also known as fission yeast), which is useful in the study of DNA replication because it has a simple genome that replicates in a manner comparable to mammalian cells (see Introduction 1.3.1). With *S. pombe* extracts Nestoras *et al.*, (2010) saw a continuum of possible R1 complex sizes using size exclusion chromatography (Figure 4.7). The majority of the product eluted at a size that seemed to correspond to an $\alpha_2\beta_2$ complex. The largest possible size might have corresponded to a $\alpha_6\beta_2$ complex. A slight increase in quantity of protein eluting at this larger size was seen as a result of treatment with HU, a chemical which inhibits RNR activity (Lammers and Follmann, 1984). Discrete populations of a particular size were not seen. The complex was found not to be stable during purification so a cross-linking treatment was used beforehand and, as such, the larger order complexes may be an artefact of cross-linking. It is therefore of great biological interest to develop a technique which can give a measure of stoichiometry *in vivo*.

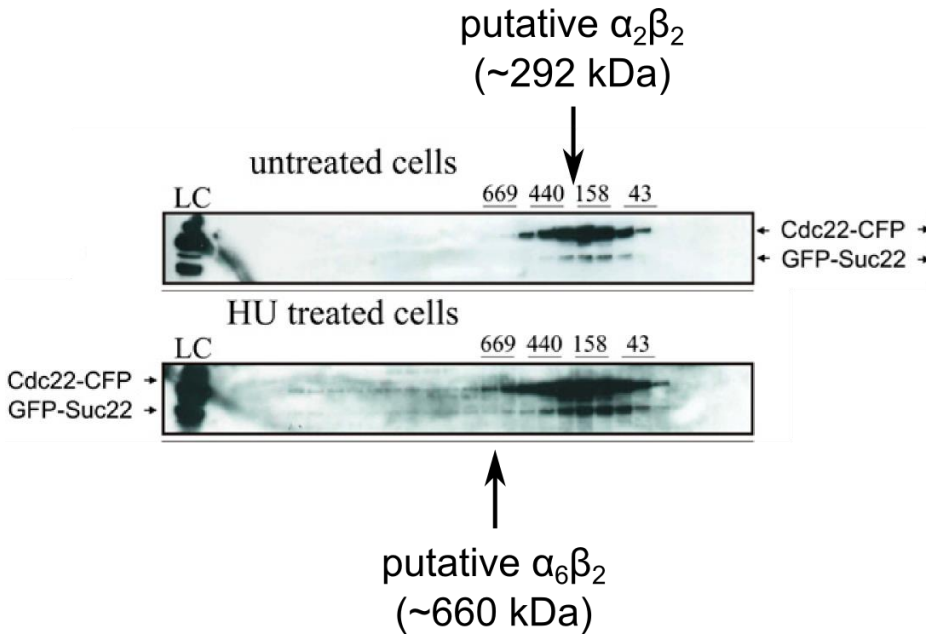


Figure 4.7 Size exclusion chromatography of the RNR complex

Figure taken from (Nestoras *et al.*, 2010)

Nestoras *et al.*, extracted the RNR complex from cells and used cross-linking and size exclusion chromatography to attempt to establish the stoichiometry of the RNR complex. They compared RNR extracted from unperturbed cells with RNR extracted from cells which had been treated with the RNR inhibitor HU. Shown above is the western blot of the different fractions obtained. An $\alpha_2\beta_2$ complex should have a molecular weight of ~292 kDa and $\alpha_6\beta_2$ ~660 kDa. In the untreated cells the majority of the protein came off at a size suggesting the $\alpha_2\beta_2$ complex, however in the HU treated extract the tail seems to be lengthened towards the larger molecular weights, including one that could represent the $\alpha_6\beta_2$ complex .

4.1.3 Objectives: Can FCS be used to observe different stoichiometries of the RNR complex?

Having established the limits of density on quantitative PALM it seemed unlikely that technique could be used to study the stoichiometry of the RNR complex *in vivo*. As such an alternative technique that measures diffusion speed was investigated to establish whether it could elucidate the size of RNR complexes. The technique FCS has been used to study diffusion speed *in vivo*. Although the $Cdc22^{R1}$ subunit is very highly expressed, it was reasoned that the photo-activatable fluorophore, mEos3.2, could be used to reduce the number of fluorophores in the FCS detection volume. The main objectives of this study were: To characterise the FCS system and analysis methods using Rhodamine B dye in solution, purified mEos3.1 in solution and mEos3.1 expressed in cells. To use the monomer and trimer constructs of mEos3.1 to establish whether a small change in size could be differentiated by FCS. To compare the conventional FCS fitting method with the maximum entropy method. To take FCS measurements with mEos3.2 tagged $Cdc22^{R1}$ *in vivo* and to compare them with measurements of cells in which HU treatment has been used to inhibit the RNR complex. It was hoped that two or more different diffusion speeds would be seen for $Cdc22^{R1}$ and that either these speeds or their relative ratios might change as a result of HU treatment.

4.2 Results

4.2.1 Using mEos to reduce the number of active fluorescent molecules in the detection volume

As has been noted, FCS experiments work best when the number *N*, the average number

of fluorescent molecules in the detection volume, is small (see 4.1.1.1). As the expression level of the protein of interest, $Cdc22^{R1}$, is very high it was necessary to find a means of reducing the number of molecules of $Cdc22^{R1}$ which were fluorescent. Looking at $Cdc22^{R1}$ N-terminally tagged with GFP (produced by Dr Adam Watson (University of Sussex) using the Cre-mediated cassette exchange system (Watson *et al.*, 2008)) it was decided that a potential approach for this was through photo-bleaching the majority of the fluorescent molecules and measuring what remained.

One means for achieving this was simply increasing the power density of the activation laser, using the neutral density filters. Illuminating with a moderate power density and leaving it for several seconds should photo-bleach molecules moving through the detection volume (Figure 4.8 A). The effect was assessed by returning to the power being used for measurement and checking the intensity trace: The general background intensity should be low with clear fluctuations showing and without a general downward trend. Care had to be taken to avoid damaging the cell so the highest possible laser powers were not used. However this technique only allowed photo-bleaching of molecules passing through the detection volume and, in the case of GFP tagged $Cdc22^{R1}$, the density was such that when the laser power was reduced insufficient molecules had been photo-bleached and new fluorescent molecules flooded back in. Little reduction in the *N* value was seen.

An alternative protocol was tested that involved flipping in a TIRF lens to de-focus the beam and turning the laser power up high in order to photo-bleach the whole cell (Figure 4.8 B). This was not completely successful as the laser used was insufficiently powerful to rapidly bleach the cell with the more diffuse beam. As an alternative to photo-bleaching for reducing the number of fluorescent molecules within the detection volume, $Cdc22^{R1}$ was N-terminally tagged with the photo-activatable fluorescent protein mEos3.2 (strain also produced by Dr Watson). Tagging with mEos3.2 allowed the use of photo-activatable FCS

(Kaur *et al.*, 2013a). A low power density of 405 nm wavelength light was used to activate a subset of the mEos3.2 fluorophores, keeping the number of activated fluorophores very low (Figure 4.8 C).

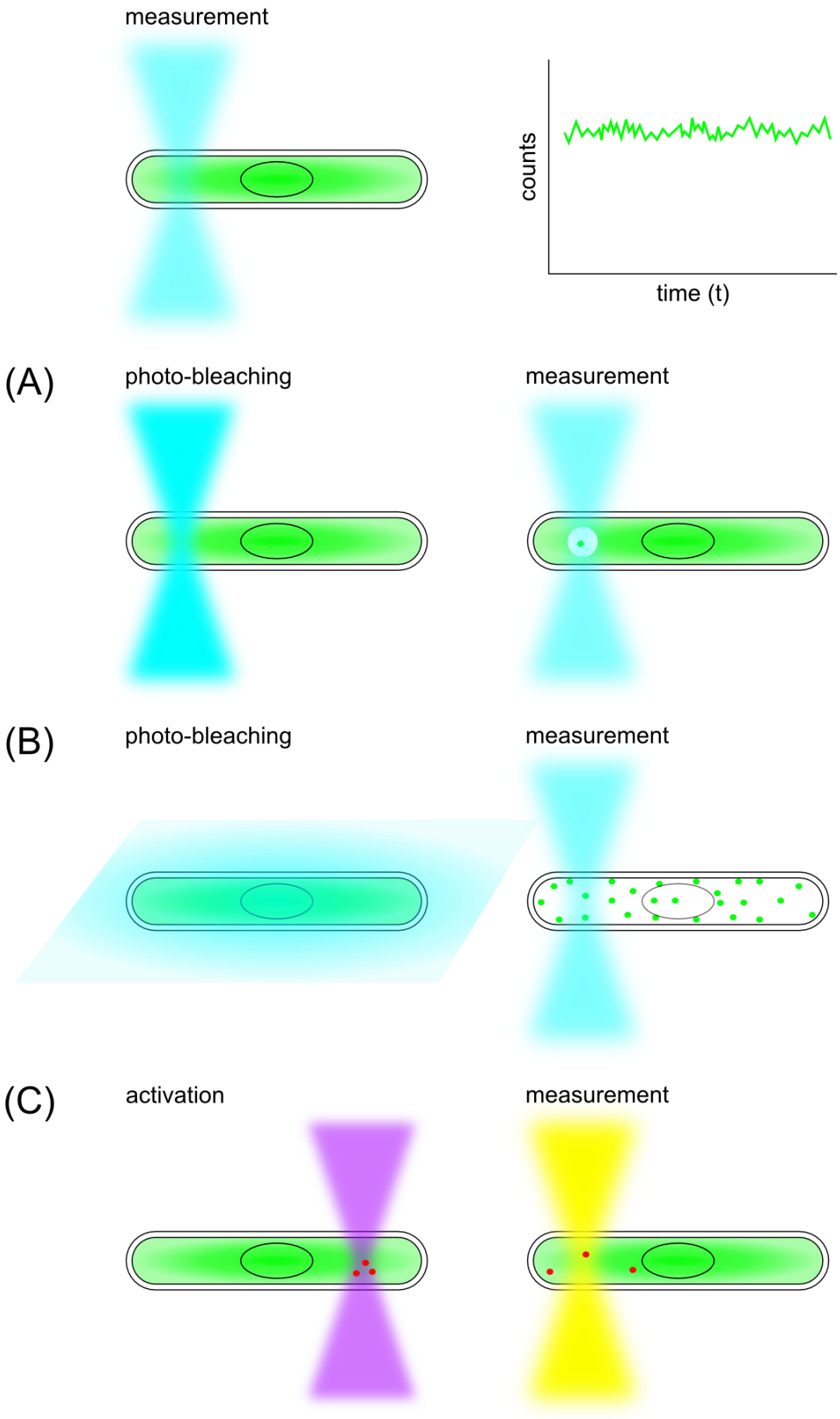
Figure 4.8 Solutions for reducing the number of fluorophores in the detection volume

$Cdc22^{R1}$ is a very highly expressed cytoplasmic protein and as a result if it is tagged with an ordinary fluorescent protein such as GFP the average number of molecules in the detection volume in an FCS experiment will be very high. For ideal FCS conditions it is best to keep this number low, such that each molecule contributes significantly to the overall signal. If this is not the case the fluctuations being measured may be lost in the noise. A few solutions have been tried to reduce this number.

(A) The power of the excitation laser was increased such that the majority of molecules passing through the area illuminated by the focused beam would be photo-bleached. This would allow measurement of a smaller number of unbleached molecules as they diffused back into the bleached area.

(B) A TIRF lens was added to the beam path on a flip hinge to allow simple switching between a focused beam and a de-focused beam. The de-focused beam was used to photo-bleach fluorescent molecules across the whole cell to reduce the number in the detection volume.

(C) The $Cdc22^{R1}$ protein was tagged with mEos3.2, a photo-activatable protein. Using laser light of 405 nm wavelength the mEos3.2 is switched from a green fluorescent form to a red fluorescent form. By using a very low power of 405 nm light only a subset of the fluorophores should be activated. A laser that excites the red form can then be used as the excitation beam.


measurement
counts
time (t)
(A)
photo-bleaching
measurement
(B)
photo-bleaching
measurement
(C)
activation
measurement

4.2.2 Measurement of the spot size and laser power density

Pixel size was calculated using images of a USAF 1951 resolution chart. A USAF 1951 resolution chart is a glass slide with a pattern of bars of standard sizes. The bars were imaged on a USB2.0 complimentary metal-oxide-semiconductor (CMOS) camera (see Materials and Methods: 2.5.1) and the images measured in pixels, then the size of the bar in μm/nm was looked up using a table online ("1951 USAF resolution test chart," 2016). This allowed the size of the pixels in the images taken by the CMOS camera to be calculated.

Using a clean coverslip, with no fluorophores, images of the reflection of the excitation beam light from the bottom of the coverslip were taken at various power levels, as controlled by the neutral density filters, to give images of a spot of light. From this the area of the spot (a 2D projection of the detection volume) was calculated, as well as its radius which is required for calculating diffusion coefficient (Table 4-1). Intensity (in arbitrary units) was plotted against pixel position to give an approximately Gaussian distribution for both the *x* (Figure 4.9) and the *y* (not shown) dimensions. From this the full width half maximum (FWHM) was used as the diameter, calculated as 2.35× the standard deviation (SD) of the intensity distribution of the spot. An overall diameter was approximated as the average of the *x* and *y* diameters (Table 4-1). The power of the 561 nm laser at the sample was measured for each neutral density filter pair using a power meter (Thorlabs). This power value was converted to a power density using the spot area (Table 4-2).

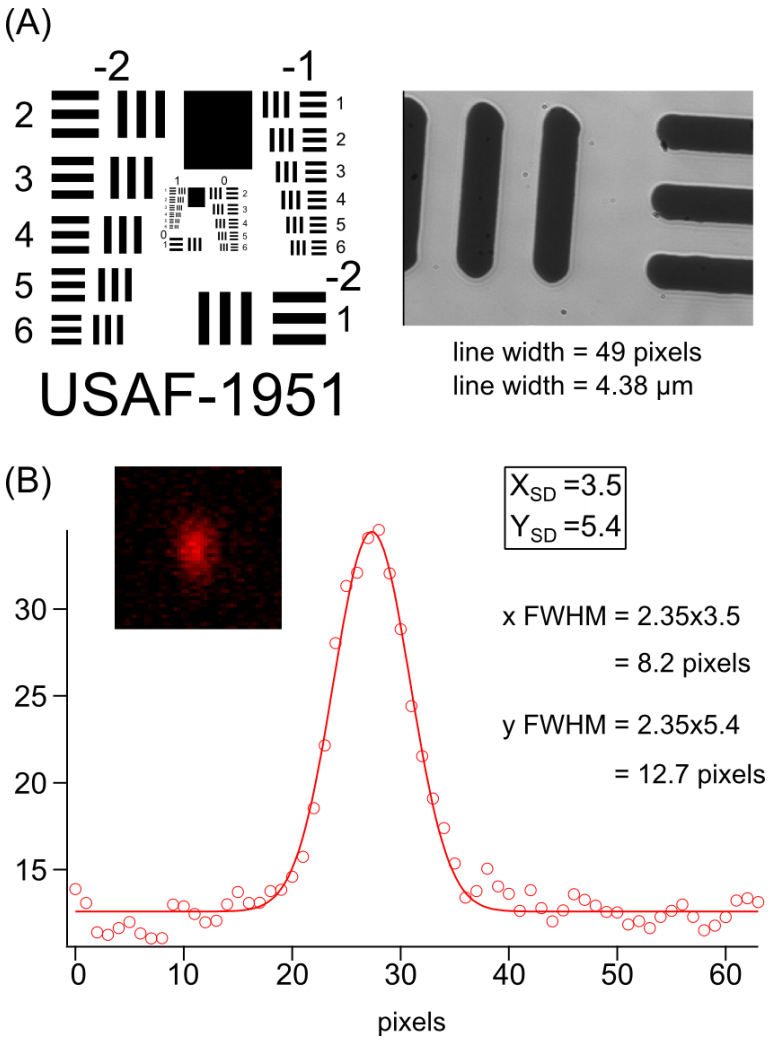


Figure 4.9 Calculating the pixel size and measuring the spot size

(A) A USAF 1951 resolution chart (example image left © User:Setreset /Wikimedia Commons/CC-BY-SA-3.0) was imaged with a CMOS camera to measure the pixel size. Bars shown (right) are size 6-6 (4.38 µm width).

(B) Images of the reflection of the excitation beam light from the bottom of the coverslip were taken using a CMOS camera to give an image of a spot (top left corner). From this the intensity of pixels across the spot in the *x* and *y* directions were measured from which the FWHM (2.35 × S.D.) is approximately the spot diameter.

X FWHM	8.2	pixel
Y FWHM	12.7	pixel
Pixel size =	89.5	nm/pixel
X	733.90	nm
Y	1136.65	nm
	0.00007339 0.00011367	cm cm
A	2.62068×10^{-8}	cm^2
diameter	935.275	nm
radius	467.638	nm

Table 4-1 Calculating spot size based on pixel size

The pixel size was calculated by imaging a USAF 1951 resolution chart (Figure 4.9). The spot size was calculated by recording the image of the reflection of light from the bottom of the coverslip. This gave an image with an approximately normal distribution in the *x* and *y* planes. The spot diameter was recorded as the width at FWHM or 2.35× the standard deviation.

ND value	power at ND filter wheel (μW)	power at sample (μW)	power density at sample (kW/cm^2)
3.5	9.0	3.8	0.1450
3.4	11.9	5.0	0.1908
3.3	15.8	6.7	0.2557
3.2	19.0	8.1	0.3091
3.0	31.0	13.6	0.5174
2.5	128.0	56.0	2.1368
2.4	173.0	73.0	2.7855
2.3	225.0	98.0	3.7395
2.2	274.0	112.0	4.2737
2.0	442.0	188.0	7.1737
1.5	1360.0	570.0	21.7501

Table 4-2 Power density at the sample for each value of neutral density filter

For various combinations of neutral density filter the power was measured at the filter wheel and at the sample. Power density was calculated by dividing the power at the sample by the calculated spot area.

4.2.3 Photo-physical effects of increasing laser power on the simple fluorophore Rhodamine B

To characterise the system FCS was performed using a 100 nm solution of Rhodamine B that has a λ_{ex} of 553 nm, meaning it can be excited by the 561 nm laser. This was used to determine whether the triplet state (a non-fluorescent excited state) was affected by the laser power density used for measurements, as this would affect the maximum entropy distributions seen. 100× 3 second acquisitions were performed at a series of different power densities that were controlled by a series of neutral density filters. The autocorrelation curves for these acquisitions were then averaged, any curves deviating significantly from the average were rejected and the average was re-calculated (see Material and Methods 2.5.2). With conventional FCS fitting the best fit was found with a single component model. With maximum entropy fitting, a single peak was seen with a narrow distribution. The diffusion time (average time a molecule spends in the detection volume) that was obtained from conventional fitting of this data and the peak values from the maximum entropy fitting had similar values. When the peaks were normalised to 1 it was possible to see that as the excitation laser power increased the peaks shifted towards a longer time component (Figure 4.10). This is probably because the effective size of the detection volume increases as laser power increases meaning that for a constant diffusion speed the diffusion time will increase with increasing power. From this experiment it can also be seen that the fractional contribution of the triplet state is lowest in the power density range of 2-8 kW/cm^2 (Figure 4.11) suggesting this would be a good power range to measure with. However, in the maximum entropy fitting the triplet state gave a peak at such a short time (~10^{-6} s) that it is unlikely to interfere with the expected peaks for the diffusion components regardless of the power used.

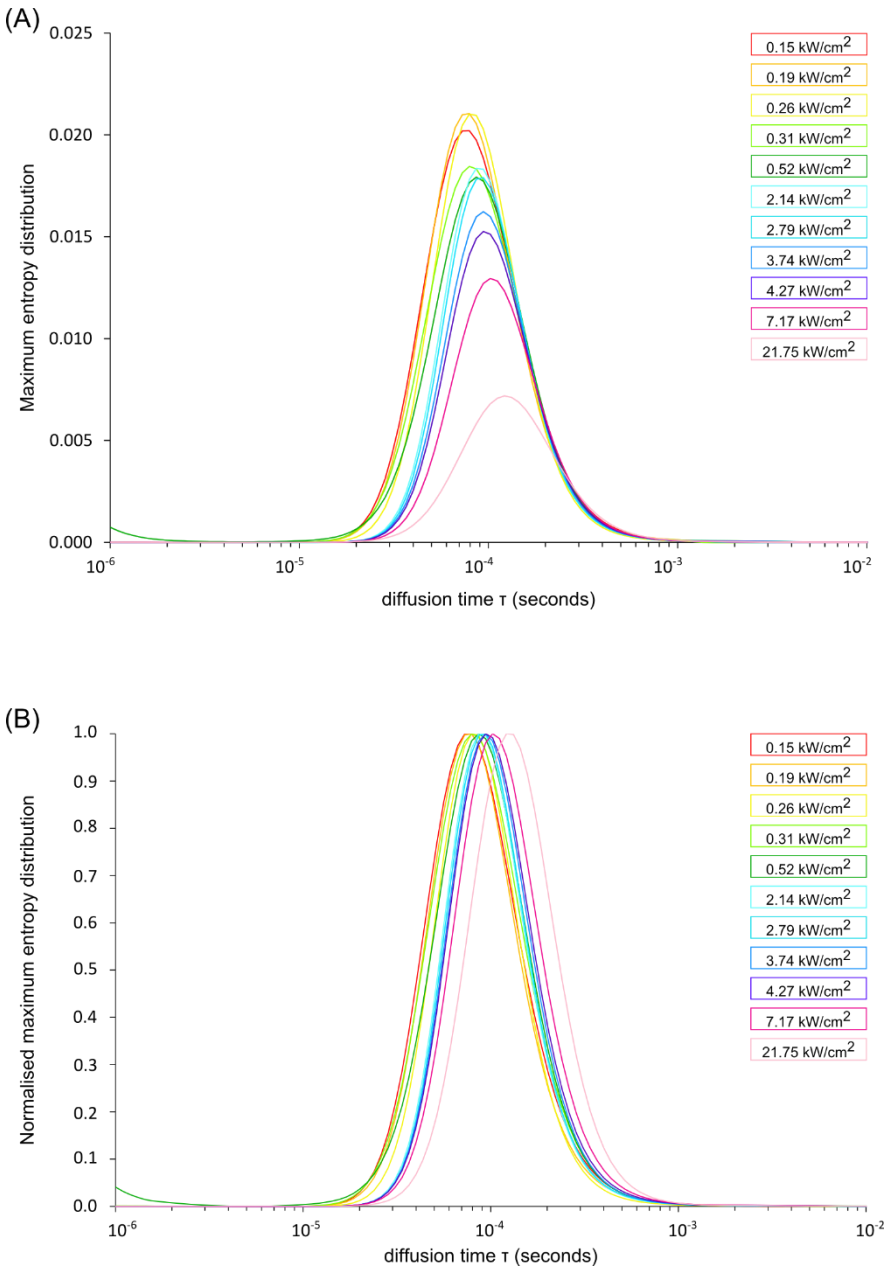


Figure 4.10 The change in maximum entropy distribution as a result of increasing laser power for Rhodamine B

(A) The maximum entropy distribution for Rhodamine B at varying power densities of the excitation laser. It is possible to see that the peaks shift to longer diffusion times with increasing power.

(B) The maximum entropy distribution with peaks normalised to 1. A clear shift is visible as the laser power is increased, with larger shifts visible where the difference between power densities is greater.

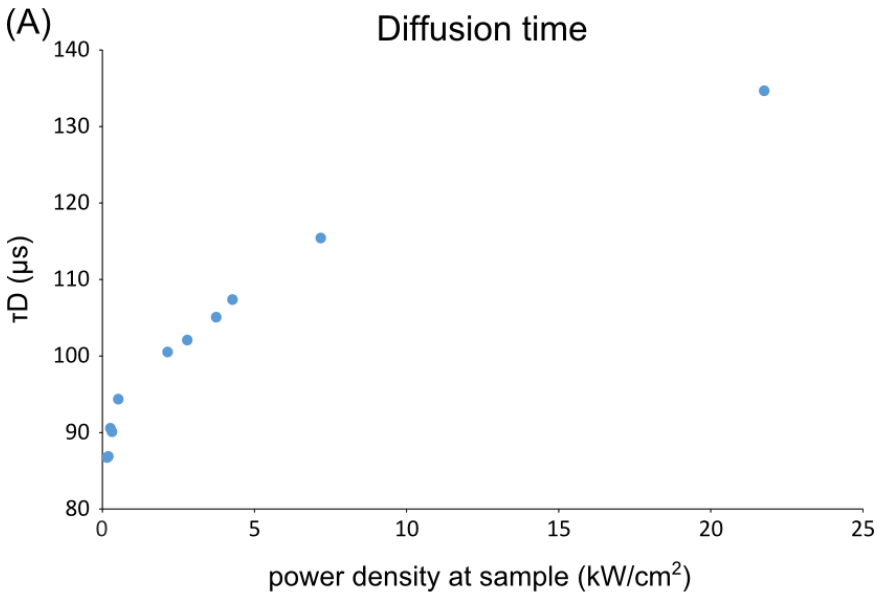
Figure 4.11 Diffusion time and triplet contribution at differing power densities from conventional fitting

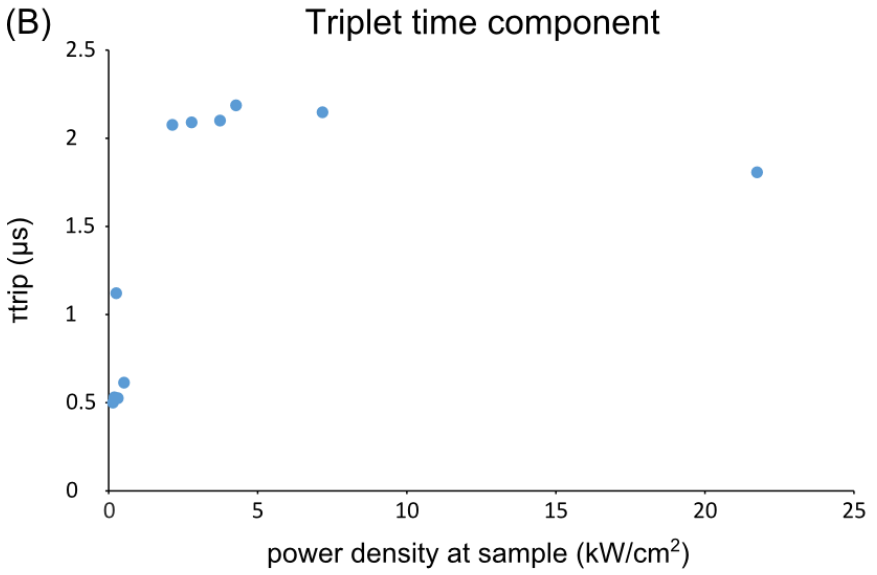
The conventional fitting plugin allows a triplet state non-fluorescent component to be included in the fitting. This triplet state has a fractional component and a time component.

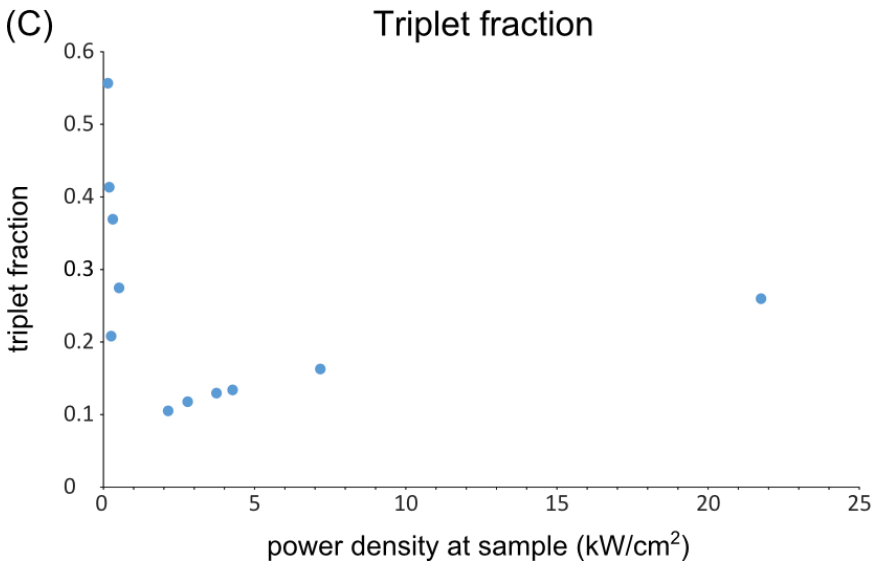
(A) The diffusion time results from conventional fitting with a single component fit. The shift to slightly longer diffusion times for greater power densities seen in the maximum entropy distribution is also seen in the result for the conventional fit.

(B) The triplet state time component from conventional fitting also shows a change with changing laser power. It appears to increase with increasing low laser powers up to ~5 kW/cm^2, it then starts to decrease at the higher laser powers (>5 kW/cm^2).

(C) The triplet fraction decreases at laser powers up to 2 kW/cm^2 then increases at powers > 2 kW/cm^2. The triplet fraction is lowest between powers 2.14 and 7.17 kW/cm^2.


(A)
Diffusion time
τD (μs)
140
130
120
110
100
90
80
0
5
10
15
20
25
power density at sample (kW/cm2)


(B)
Triplet time component
τtrip (μs)
2.5
2
1.5
1
0.5
0
0
5
10
15
20
25
power density at sample (kW/cm2)


(C)
Triplet fraction
triplet fraction
0.6
0.5
0.4
0.3
0.2
0.1
0
0
5
10
15
20
25
power density at sample (kW/cm2)

4.2.4 Photo-physical effects of increasing laser power on free mEos3.1 in solution

Having established how the values from conventional fitting and maximum entropy change for Rhodamine B dye, the same power dependency experiment was reproduced using purified mEos3.1 in solution, as this was the fluorophore to be used *in vivo*. As a fluorophore mEos3.1 is dimmer than Rhodamine B but has the benefit that photo-activation allows artificial control of the concentration of active fluorophore during the experiment. In contrast to the Rhodamine B , the mEos3.1 curves fitted better with the 2 component conventional fit. With the maximum entropy fit, at lower powers a single peak was seen, at higher powers this split into two peaks, with the shorter time peak shifting leftwards (decreasing) (Figure 4.12). This gave a short time peak with an unfeasibly fast diffusion coefficient (~200 $\mu m^2/s$ shifting to ~1300 $\mu m^2/s$) and a long time peak that gives a diffusion coefficient (30-45 $\mu m^2/s$) that is more in range with the reported diffusion of GFP in water (87 $\mu m^2/s$) (Elowitz *et al.*, 1999). The short time component possibly represents a photo-physical effect such as photo-bleaching or blinking. One possible explanation is that at low powers this photo-physical effect has a maximum entropy distribution which overlaps with the diffusion distribution and thus a single, broad peak is seen. Whereas, high laser powers encourage the photo-physical effect to happen more rapidly, shifting the peak to the right of the graph and revealing the diffusion peak, which is possibly also shifting leftwards as seen in the Rhodamine B experiment.

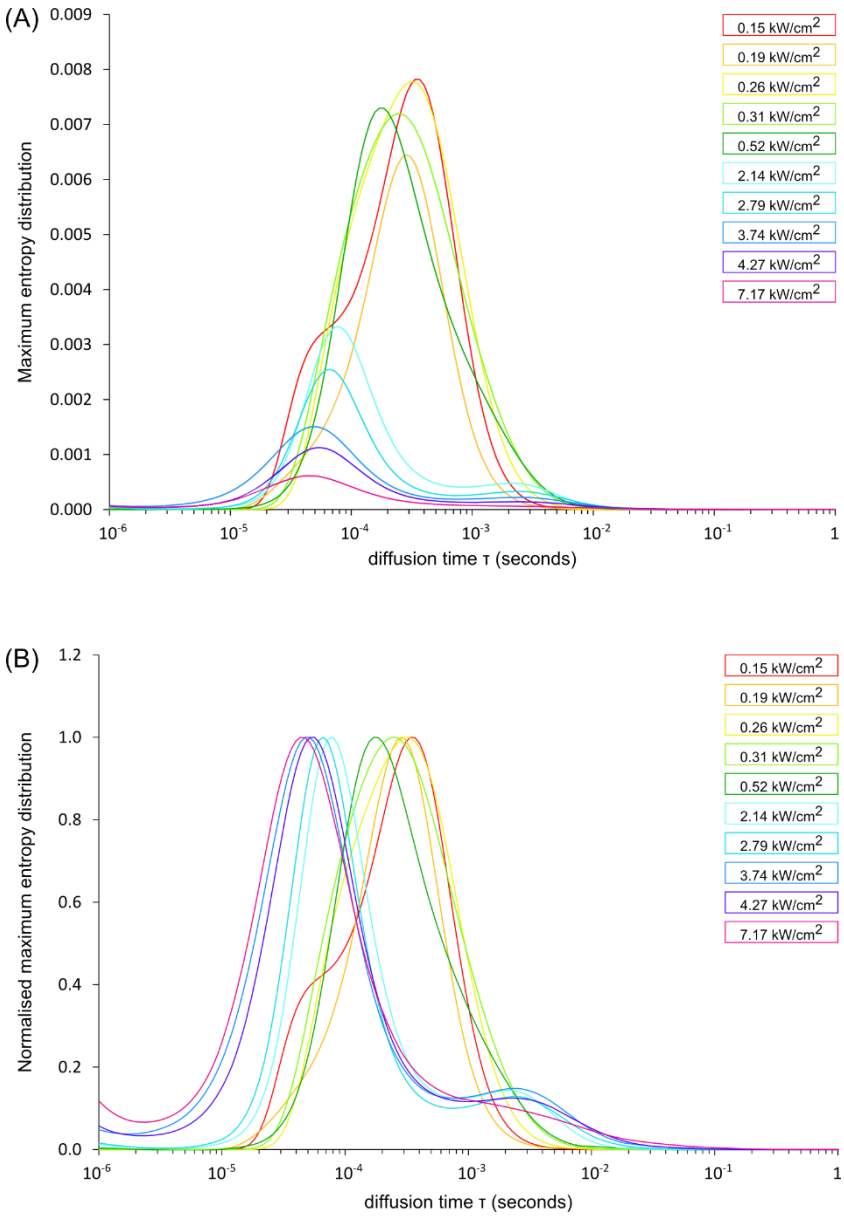


Figure 4.12 The effect of increasing laser power on freely diffusing mEos3.1 in solution in H_2O

(A) Using maximum entropy fitting at low powers a single, broad peak was plotted. At increasing laser powers the peak splits into two parts. The peak at the slower time is most likely to represent the diffusion time of mEos3.1, giving a diffusion coefficient of 30-45 $\mu m^2/s$.

(B) The maximum entropy distributions were normalised to their maximum values such that the highest peak for each is equal to 1.

4.2.5 Photo-physical effects of increasing laser power on free mEos3.1 in the cytoplasm of the cell

Having characterised the FCS system *in vitro* it was important to compare the results to a similarly simple system *in vivo*. As such, FCS was performed in cells expressing mEos3.1 as a cytoplasmic protein. FCS performed inside a cell differs from FCS performed in a solution in two important ways: one is that the cellular environment has more noise, with auto-fluorescent compounds existing in the cell as well as the fluorescent protein being measured. The other is that the cytoplasm of the cell is a crowded fluid, having in solution a great many particles of varying sizes. As was observed in early *in vivo* FCS experiments, diffusion coefficients in cells are slower than those in water due to the increased viscosity this crowding creates (Berland *et al.*, 1995; Politz *et al.*, 1998). In addition, molecules diffusing in crowded fluids can undergo what is known as anomalous diffusion (reviewed by Höfling and Franosch, 2013), and the scale of the effect of anomalous diffusion varies depending on the size of the observed molecule (Weiss *et al.*, 2004).

FCS acquisitions were performed in cells expressing mEos3.1 in the cytoplasm under the control of *urg1*, the uracil-inducible promoter, and the *spo5* DSR - the transcript removal element (see Chapter 3: 3.2.1). With cellular mEos3.1, the same separation of two peaks that was noted for mEos3.1 in solution was observed (Figure 4.13). This seemed to mainly be from the left-most peak shifting left: the right-most peak stayed more or less in the same place. With mEos3.1 in solution, the peaks at the lowest powers appeared to overlap. However, in cells two separate peaks were observed at all powers. At very low powers (0.15 and 0.19 kW/cm^2) the distribution differed from the distribution at higher powers, it is likely that this was due to having a poor SNR at these lower laser powers. As was suggested for the mEos3.1 in solution it is likely that the right-most peak is the one that represents the diffusion time of mEos3.1. The right-most peak was at a longer time in this cellular

experiment than that seen for mEos3.1 in solution and this is what would be expected as a result of the crowded environment of the cytoplasm. Interestingly, the left-most peak is in almost the same place, again leading to the conclusion that it represents a photo-physical component, which would be unaffected by changes in viscosity, rather than a diffusional one. One conclusion to draw from this is that the very low powers are not the most suitable for carrying out these experiments. Conventional fitting gives a slow diffusion coefficient of 10-15 $\mu m^2/s$, the diffusion coefficient of GFP in *E. coli* has been reported as 4-8 $\mu m^2/s$ (Elowitz *et al.*, 1999). The difference between GFP and mEos3.1 may be due to a difference in the dynamic radius of mEos3.1 compared to GFP, or to a difference in method of calculating spot size or diffusion coefficient, or to a difference in the viscosity of *S. pombe* in comparison to *E. coli*, the fact that they are in the same order of magnitude is positive.

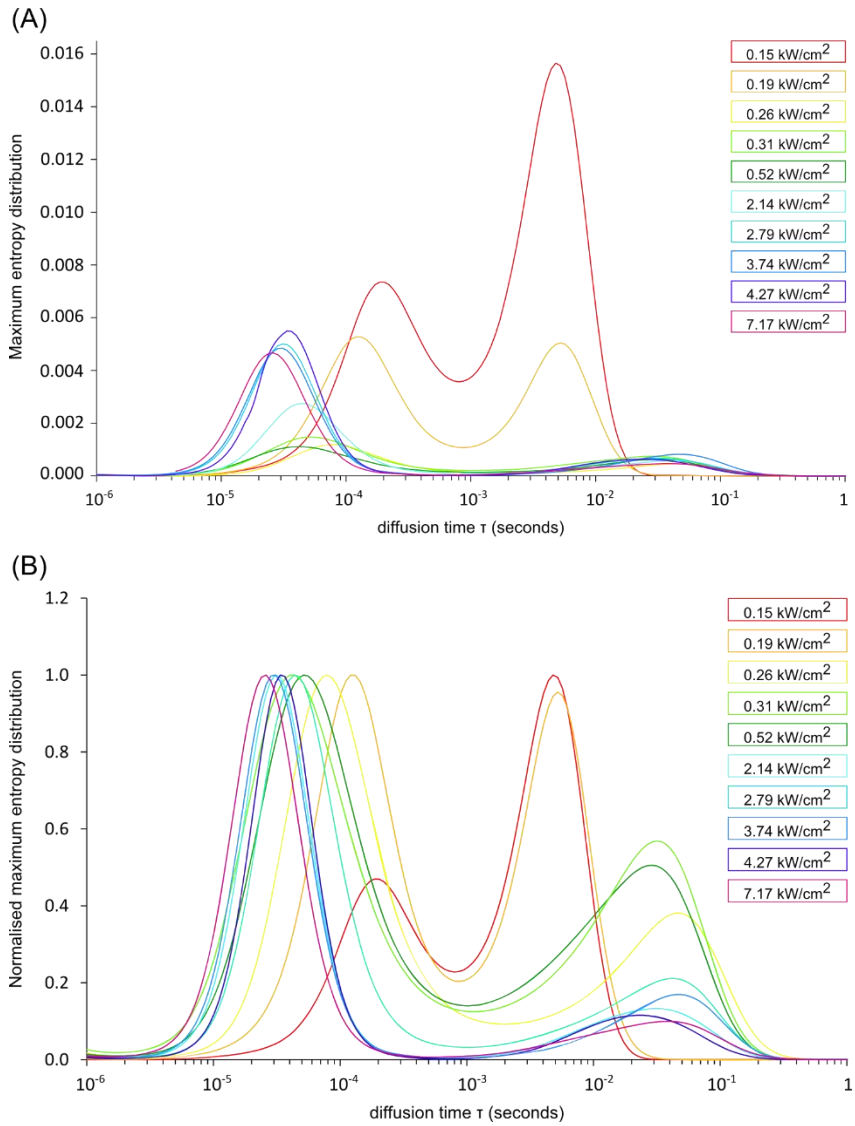


Figure 4.13 The effect of increasing laser power on mEos3.1 diffusing freely in a cell

(A) Maximum entropy distribution for a series of measurements taken using cells expressing mEos3.1 in the cytoplasm. Two peaks were fitted for cellular mEos3.1 at all laser powers. The distribution for the lowest two powers differs noticeably from the rest of the measurements.

(B) When the data are normalised to the value of the highest peak it becomes clear that there is the same leftward shift in the left-most peak that was seen with mEos3.1 in solution. The right-most peak does not appear to be shifting in the same way in either direction.

4.2.6 Using FCS it is possible to differentiate between a monomer and a trimer of mEos3.1

Having characterised how the fluorophore mEos3.1 behaved both in solution and in the cytoplasm of a cell the next question to ask was whether a difference in diffusion time could be seen between molecules with only a small difference in molecular weight (MW). Using an excitation laser power of 7.17 kW/cm^2 measurements were taken in ~10 cells each of strains expressing either a monomer or a trimer of mEos3.1 (as described in Chapter 3: 3.1.3). The diffusion coefficient of freely diffusing molecules is proportional to their hydrodynamic radius, which scales with the cube root of molecular weight, assuming a roughly spherical protein (Haustein and Schwille, 2007). As such, it was uncertain whether it would be possible to see a difference between the monomer and the trimer, as a difference of 3 fold in MW would only lead to a 1.44 fold change in diffusion speed. However, two factors might emphasise the difference: first, a trimer of mEos3.1 is not a single globular protein, but three structures in a short "string". Second, there is the possibility that in a cellular environment, due to molecular crowding, differences in hydrodynamic radius would effectively be amplified as large molecules would be more slowed by the crowding than smaller ones.

For each cell, 100× 3 second measurements were taken and averaged in order to avoid giving undue weight to sample artefacts (Ries *et al.*, 2010). The average curves were fitted with a two component equation and a maximum entropy distribution (Sengupta *et al.*, 2003). The maximum entropy distributions for all of the cells were normalised such that the right-most peak was equal to 1 and plotted onto a single chart for comparison (Figure 4.14). The right-most peaks showed a visible separation with the peaks for trimer (blue) occurring at a longer diffusion time than the monomer (red). This means it is possible to distinguish between the diffusion times of monomer and the trimer, even though the

difference in molecular weight is quite small.

Following this, the same measurements were performed in cells expressing a dimer of mEos3.1 to see whether it was possible to distinguish molecules with only a 2 fold change in molecular weight. When the maximum entropy distributions were plotted the right-most peak for the dimer experiments did appear to sit in between the peaks for the monomer and the trimer (albeit with an overlap) and the left-most peaks all sat at more or less the same place (Figure 4.14). However, if the long diffusion times from conventional FCS fitting are plotted as a box plot the difference between trimer and monomer is clear whereas the monomer and dimer have a lot of overlap. Further, using a Welch two-sample t-test the difference in diffusion times comparing monomer to trimer gave a $p<0.05$ whereas the difference between monomer and dimer was not significant.

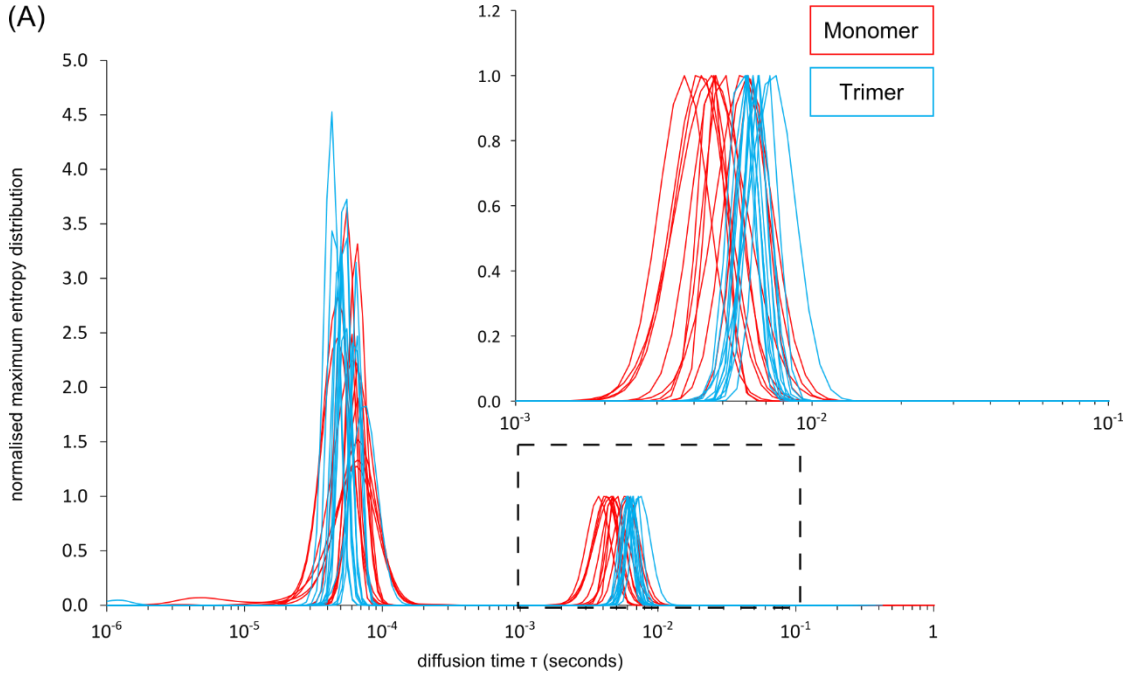
Bearing in mind that at 27 kDa the mEos3.1 and 3.2 proteins are relatively small and that the slowing caused by crowding is likely to have a greater effect the larger the molecules being observed are, this bodes well for the ability of maximum entropy fitting to distinguish between a dimeric and a hexameric complex of $Cdc22^{R1}$ *in vivo*. It also suggests that the maximum entropy fitting routine might distinguish between different, similarly sized components better than conventional fitting.

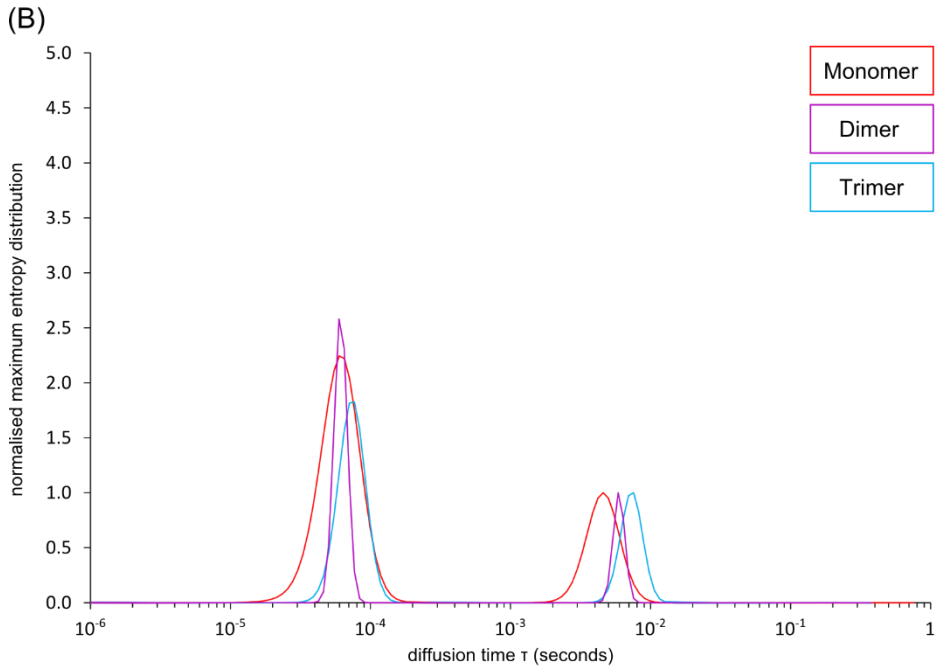
Figure 4.14 Using FCS it is possible to distinguish between a monomer, dimer and trimer of mEos3.1

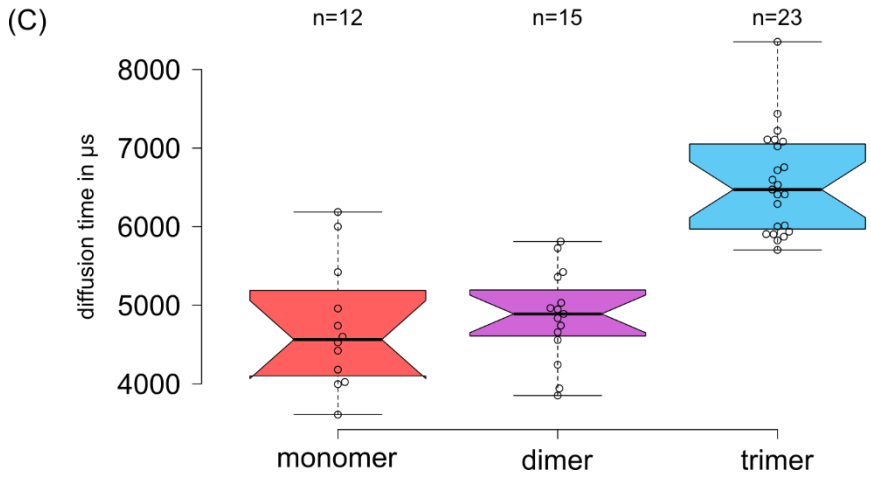
(A) Plotting the maximum entropy distributions, normalised to the right-most peak, for all of the results from cells expressing an mEos3.1 monomer (red) or an mEos3.1 trimer (blue). The left-most peak is in the same region for all of the samples. The right-most peak appears to separate into roughly two populations, monomer and trimer.

(B) Cells expressing a dimer of mEos3.1 were subsequently measured (purple). Data displayed has been normalised to the right-most peak. Taking a representative sample from each experiment it can be seen that the right-most peak for the dimer distribution falls between the monomer and trimer peaks.

(C) Taking the diffusion times calculated for the long time components using conventional FCS fitting and plotting them as a box plot the difference between monomer and trimer is much clearer than the difference between monomer and dimer.


(A)
Monomer
Trimer
normalised maximum entropy distribution
diffusion time τ (seconds)


(B)
Monomer
Dimer
Trimer
normalised maximum entropy distribution
diffusion time τ (seconds)


(C)
n=12
n=15
n=23
diffusion time in µs
monomer
dimer
trimer

4.2.7 Using FCS it is difficult to differentiate between two different sized dextrans in a mixture

With the monomer and trimer experiment it was shown that it was possible to differentiate between two molecules quite close in size. However the monomers and trimers were expressed in separate cells. It was important to test whether it was possible to distinguish two species if they were together in a mixture, as the different oligomers of $Cdc22^{R1}$ would be, and what effect the proportions of the mixture would have. Previous studies on the resolution of FCS have suggested that a difference in diffusion speed of at least 1.6 fold is required (Meseth *et al.*, 1999). To examine this in the FCS system being used Rhodamine B labelled dextrans of sizes 3kDa and 10 kDa (ChemQuest) were purchased, having a roughly 3 fold difference in MW analogous to the difference between a dimer and a hexamer of $Cdc22^{R1}$. A 20 nM stock solution of each dextran was used.

To establish whether it was possible to see a size difference to begin with, the 3 kDa and 10 kDa dextrans were measured separately at concentrations of 10 nM. Looking at the maximum entropy distribution (Figure 4.15 A) it is just possible to see a difference between the two sets of peaks. Using a conventional fit, and plotting the resulting diffusion times as a box plot, there appears to be a significant difference. However, when a 50:50 solution of each dextran at 10 nM was measured the maximum entropy distribution gave one broad peak shifted slightly to the right of either of the dextrans alone (Figure 4.15 B). If a two component conventional FCS fit is used to try to force two answers, two diffusion times are returned, one lower and one higher than the times calculated for either the 3 kDa time or the 10 kDa dextran when not in a mixture. This suggests that it will be difficult to distinguish two species in a mixture unless they differ by a greater amount. The shift to the right in the mixture may be because, although each individual dextran is at the same concentration as in the separate experiments, the solution is twice as crowded.

Mixtures consisting of 75% 3 kDa + 25% 10 kDa and 75% 10 kDA + 25% 3 kDa were also observed. There was a small shift in the maximum entropy peaks, with the solution containing more of the 10 kDa having a peak shifted to longer times (Figure 4.15 C). Using conventional fitting with a two component fit, two diffusion times were produced that did not match the diffusion times for the 3 kDa or the 10 kDa alone. However, interestingly, the proportion of the long time component versus the short time component did shift, with a greater proportion of the long time in the 75% 10 kDa mixture. The proportions given by the conventional fit did not match the proportions present in the sample, being approximately 55:45.

As mentioned above it was expected that diffusion in the cellular environment would differ from diffusion in water as a result of molecular crowding. It was hypothesised that this molecular crowding would have a greater slowing effect on larger molecules than on small ones (as in a gel matrix). To test this hypothesis the 50:50 proportion experiment was repeated with 25% polyethylene glycol (PEG), a molecular spacer, added to the solution. This had a strong slowing effect as a rightward shift was seen in the peak on the maximum entropy distribution and both diffusion times output by the conventional fit were longer (Figure 4.15 D). However, it did not cause the single maximum entropy peak to split into two, or even to broaden, so it was not sufficient to distinguish between the two different populations.

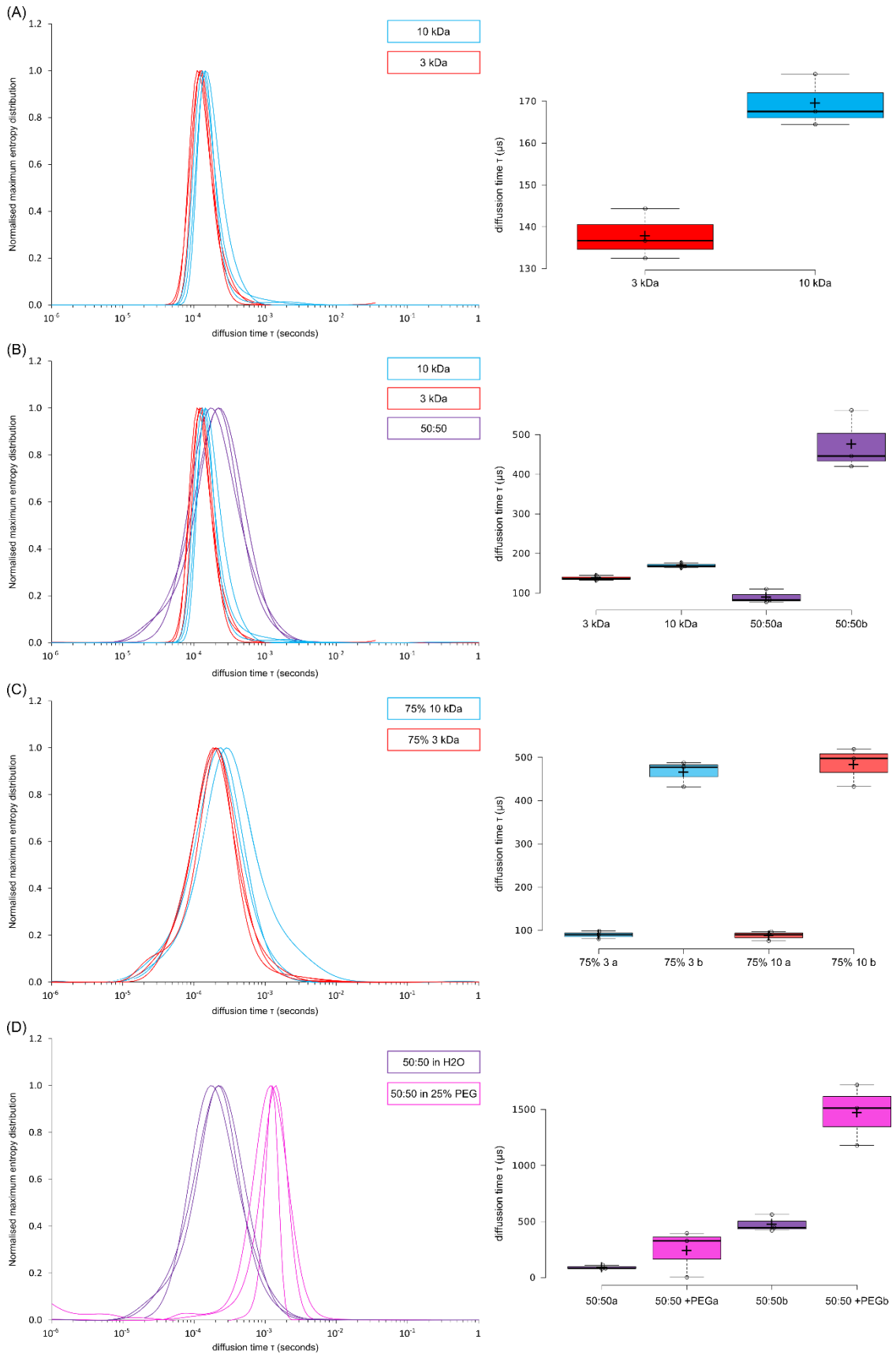
Figure 4.15 Looking at different sized fluorescently labelled dextrans in solution

(A) Rhodamine B labelled dextrans of 10 kDa (blue) and 3 kDa (red) in solution were measured separately using FCS. In both the maximum entropy distribution and the single component conventional fit it is possible to see a difference between the two.

(B) When the two dextrans were mixed in a 50:50 ratio the maximum entropy distribution showed a broader peak shifted slightly to the right of the 10 kDa peak. Fitted with a conventional two component FCS fit it gave two diffusion times, one much larger and one much smaller than the times seen individually for 10 kDa or 3 kDa.

(C) Using a 75:25 ratio of 10:3 kDa (blue) or 3:10 kDa (red) yielded a similar result to the 50:50 ratio. In the maximum entropy distribution the peak is broader than for the unmixed samples, however there is a slight rightwards shift for the mixture containing a greater proportion of the larger dextran. Using the two component conventional fitting, again the values output are much higher and lower than the values seen in the unmixed samples.

(D) The 50:50 mixture experiment was repeated in 25% PEG to study the effect of a more crowded or viscous system. Both the maximum entropy peaks and the conventional fit diffusion times were shifted to longer values in the PEG and the peaks did not broaden or split.


(A)
10 kDa
3 kDa
Normalised maximum entropy distribution
diffusion time τ (seconds)
diffusion time τ (μs)
3 kDa
10 kDa
(B)
10 kDa
3 kDa
50:50
Normalised maximum entropy distribution
diffusion time τ (seconds)
diffusion time τ (μs)
3 kDa
10 kDa
50:50a
50:50b
(C)
75% 10 kDa
75% 3 kDa
Normalised maximum entropy distribution
diffusion time τ (seconds)
diffusion time τ (μs)
75% 3 a
75% 3 b
75% 10 a
75% 10 b
(D)
50:50 in H2O
50:50 in 25% PEG
Normalised maximum entropy distribution
diffusion time τ (seconds)
diffusion time τ (μs)
50:50a
50:50 +PEGa
50:50b
50:50 +PEGb

4.2.8 Looking at Cdc22^{R1}-meos3.2 containing cells

Cdc22^{R1} is the large subunit of the RNR complex and this complex has been shown to exist in multiple stoichiometries in other organisms (see 4.1.2.1). The RNR complex is very unstable outside of the cellular environment and as such, in order to investigate its stoichiometric states, it is necessary to study it *in vivo*. While the characterisation of FCS *in vitro* and using mEos3.1 *in vivo* did not show unambiguously that we should expect to be able to distinguish the different RNR stoichiometries, it suggested that FCS may provide some information of interest as variation in diffusion rate would be related to variations in complex size and shape, and would be influenced by anomalous diffusion. Initially the Cdc22^{R1} subunit had been tagged with GFP. However, Cdc22^{R1} is very highly expressed (Carpy *et al.*, 2014) and FCS requires that a very low number of active fluorophores be in the detection volume at any one time. Thus it was decided to make use of the photo-activatable properties of the fluorescent protein mEos3.2. By activating only small numbers of the mEos3.2 fluorophores it would be possible to keep the number of fluorescent molecules in the detection volume small enough to obtain a good FCS signal (Figure 4.8).

Unfortunately, Cdc22^{R1} is an essential gene and fusion with the mEos family of fluorophores has consequences for function: anecdotally in the Carr lab multiple proteins that can be tagged with GFP without significant phenotypic consequences have not been successfully tagged with mEos fluorophores (i.e. strains with mEos family tags show significant phenotypic changes). This was also the case with Cdc22^{R1}, which had been successfully C-terminally tagged with GFP. Attempts to tag it with mEos3.2 using both the C-terminal tagging cassette exchange system (Watson *et al.*, 2008) and the Bähler system (Bähler *et al.*, 1998) lead to strains that had become diploid (data not shown). These strains contained one copy of the tagged gene and one untagged copy, suggesting that the

fluorophore was interfering with an essential function of the protein.

Dr Watson was able to create a Cdc22^{R1} N-terminal mEos3.2 strain using cassette exchange. However, after a time it was noted that cell cultures contained a proportion of cells with no fluorescence and that the number of these cells increased over time. This phenomenon has occasionally been seen before in the lab and suggests that there is selective pressure favouring the loss of mEos3.2. After propagation of cells it was often seen that there were very few that still displayed fluorescence. Lacking a sufficient number of cells that contained the fluorophore made it almost impossible to take FCS measurements in the original cultures generated by Dr Watson. Furthermore, it could not be assumed that, in the cells which did show fluorescence, a) the mEos3.2 was still attached to Cdc22^{R1} and b) that the tagged Cdc22^{R1} would behave in the normal way.

4.2.8.1 The majority of mEos3.2 is cleaved from Cdc22^{R1}

As the base strain for the cassette exchange was available the cassette exchange was repeated and two strains were isolated that showed the expected levels of mEos3.2 activation. FCS was performed on these cells as soon as they were confirmed to contain mEos3.2. The resulting maximum entropy peaks overlapped with the same area peaks seen for the monomer of mEos3.1 construct (Figure 4.16). This suggested that the mEos3.2 was being cleaved from the Cdc22^{R1} and that it was mostly free-mEos3.2 being observed. Using conventional fitting a diffusion coefficient of 8-15 $\mu m^2/s$ was obtained, similar to the one seen for free mEos3.1 in cells.

This was explored by western blot analysis. The selected cells were grown, lysed and protein extracts prepared for western blotting. Samples of cells expressing the mEos3.1 monomer and AW310 cells in which Cdc22^{R1} is not tagged were also prepared as controls. Two western blots containing the same set of samples were run and the membranes treated

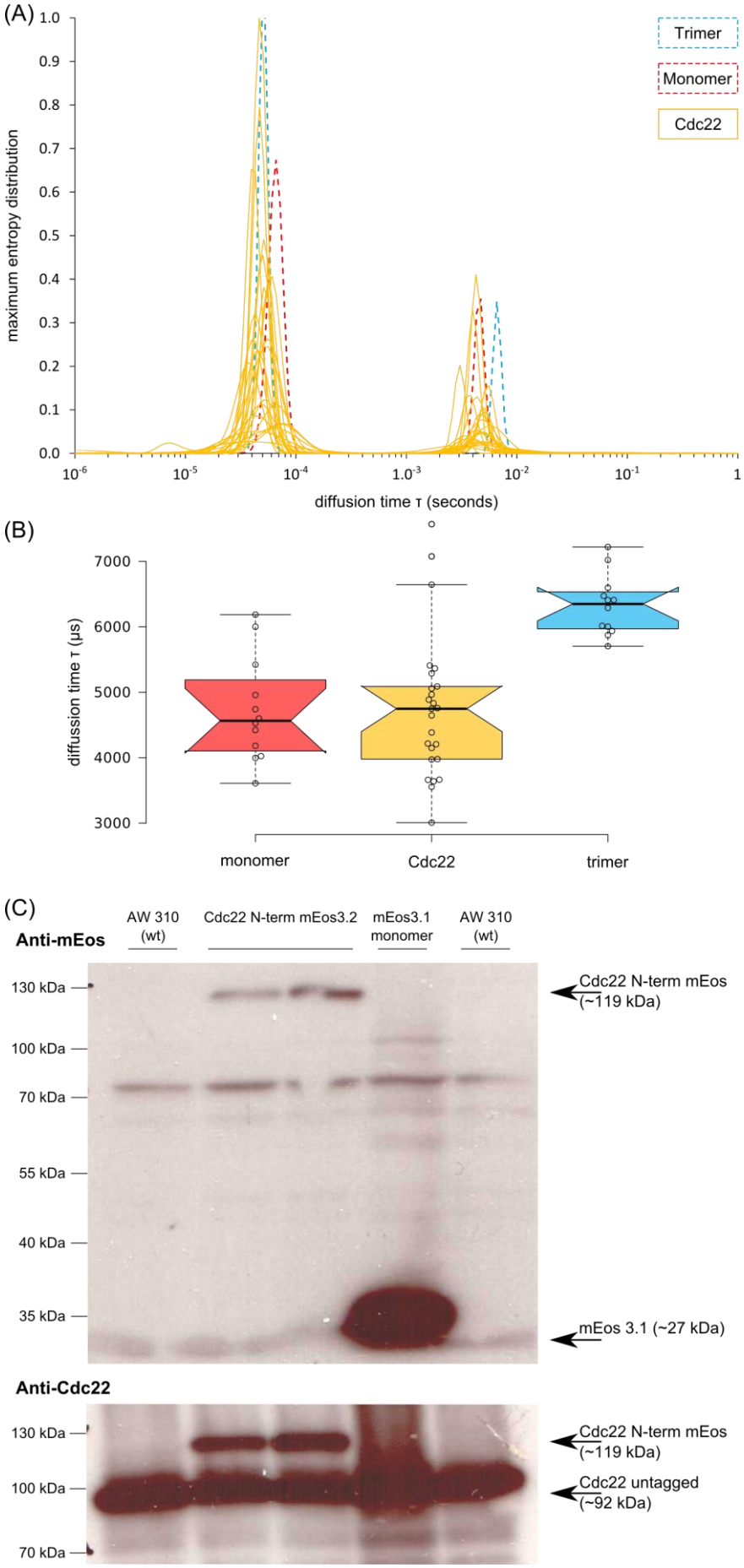
either with an antibody against the mEos protein family or an antibody against Cdc22^{R1}. With the mEos family antibody, a band can be seen in the Cdc22^{R1} mEos3.2 samples at ~119 kDa (corresponding to Cdc22^{R1} tagged with mEos3.2). A monomeric mEos3.1 can be seen at ~27 kDa in the strain expressing the monomeric mEos3.1 construct. Having measured a diffusion speed equivalent to mEos3.1 monomer in Cdc22^{R1} mEos3.2 cells, it was expected that a band at a similar molecular weight to the monomer would be seen in the Cdc22^{R1} mEos3.2 strains corresponding to the mEos3.2 that had been cleaved. However no such band was observed. The anti-mEos antibody used was raised by the Carr lab as a polyclonal antibody, so it is possible (but unlikely) that the epitope recognised by the antibody is not present on the cleaved mEos3.2, even though it would be expected to be functional and thus encompass the vast majority of the amino acid sequence. Using the anti-Cdc22^{R1} antibody a band at ~92 kDa is seen in all of the samples (corresponding to untagged Cdc22^{R1}). In the two tagged strains a fainter band at ~119 kDa (corresponding to mEos3.2 tagged Cdc22^{R1}) was also observed suggesting that there is more of the untagged Cdc22^{R1} than the tagged. (Figure 4.16), Dr Watson also observed this cleavage in the N-terminal GFP tagged Cdc22^{R1} strain (data not shown).

Figure 4.16 The mEos3.2 tag is cleaved from Cdc22^{R1}

(A) The maximum entropy distributions from a selection of cells expressing Cdc22^{R1} N-terminally tagged with mEos3.2 compared to the maximum entropy distribution for a representative sample of monomer and trimer. All distributions show the peak just below 1×10^{-4} s. The right-most peaks for the Cdc22^{R1} distributions overlap well with the peak for the monomer.

(B) Using conventional fitting to compare the long diffusion times from the Cdc22^{R1} data with the monomer and trimer data the values are very similar to those for monomeric mEos3.1.

(C) Western blots were performed using a wild type strain in which Cdc22^{R1} is not tagged, two Cdc22^{R1} N-term mEos3.2 strains and a strain expressing the mEos3.1 monomer. Two membranes were produced, the membranes were either treated with an antibody to the mEos protein family or an antibody to Cdc22^{R1}. With the antibody to mEos some mEos3.2 tagged Cdc22^{R1} can be seen at ~119 kDa and monomeric mEos3.1 can be seen at ~27 kDa in the monomer strain. Any bands which appear in all samples including the wildtype can be assumed to be non-specific. It was expected that a band at a similar molecular weight to the monomer would be seen in the Cdc22^{R1} strains corresponding to the mEos3.2 that had been cleaved. However, no such band was observed. Using the Cdc22^{R1} antibody a band at ~92 kDa is seen in all of the samples (corresponding to untagged Cdc22^{R1}). A fainter band in the tagged strains at ~119 kDa (corresponding to mEos3.2 tagged Cdc22^{R1}) was also observed.


(A)
maximum entropy distribution
diffusion time τ (seconds)
Trimer
Monomer
Cdc22
(B)
diffussion time τ (µs)
monomer
Cdc22
trimer
(C)
Anti-mEos
AW 310 (wt)
Cdc22 N-term mEos3.2
mEos3.1 monomer
AW 310 (wt)
130 kDa
100 kDa
70 kDa
55 kDa
40 kDa
35 kDa
Cdc22 N-term mEos (~119 kDa)
mEos 3.1 (~27 kDa)
Anti-Cdc22
130 kDa
100 kDa
70 kDa
Cdc22 N-term mEos (~119 kDa)
Cdc22 untagged (~92 kDa)

4.2.8.2 There is a third peak at a long diffusion time which may correspond to an oligomer of $Cdc22^{R1}$

When the maximum entropy distributions were normalised (to either peak) a third peak at very long diffusion times became visible in some of the distributions (Figure 4.17 A). Looking specifically at an FCS curve that produced a three peak maximum entropy distribution, fitting was attempted with a two or a three component conventional fit to extract a diffusion time to match the third maximum entropy peak. When this was done it could be seen that the fitting program tends to give an additional faster component rather than finding an equivalent slower component (Figure 4.17 B). Possibly the contribution from this third peak is too low for conventional fitting to find it. In any event, evidence for a very slow component can only be observed with the maximum entropy fitting method.

Since the part of the maximum entropy distribution that contains the peaks of potential interest exists above 5×10^{-4} it was decided to re-analyse only that part of the FCS curves. When this was done the maximum entropy distribution peak that overlapped the mEos3.1 monomer was shifted to a shorter time and various peaks at longer times were revealed (Figure 4.17).

The $Cdc22^{R1}$ subunit has a MW of ~92 kDa, the $Suc22^{R2}$ subunit is ~54 kDa and mEos3.1/2 is ~27 kDa, this means that a complex of $\alpha_2\beta_2$(+2mEos) would have a MW of ~346 kDa and $\alpha_6\beta_2$(+6mEos3.2) would have a MW of ~822 kDa. There would be a roughly 30 fold difference in MW between the $\alpha_6\beta_2$ and cleaved mEos3.2 and 12 fold between $\alpha_2\beta_2$ and cleaved mEso3.2. This should give diffusion speed differences of 3.1 and 2.3 fold (assuming a roughly spherical protein). The suggested resolution limit of FCS is a 1.6 fold difference in diffusion speed (Meseth *et al.*, 1999) suggesting it should be possible to observe differences between cleaved mEos and either of the hypothetical complexes. The difference between an $\alpha_2\beta_2$ and an $\alpha_6\beta_2$ complex would be 2.4 fold in MW giving

approximately a 1.33 fold difference in diffusion speed. The results from maximum entropy fitting of a monomer, dimer and trimer of meos3.2 suggest this difference may be resolvable. However, the results from using a mixture of rhodaminated dextrans suggest that differentiating components in a mixture is more difficult than when they are measured separately. Therefore, it is probably not going to be possible to use FCS to prove that the RNR complex exists in these specific stoichiometries. However, it should still be possible to investigate whether the maximum entropy peaks shift or change shape under conditions that affect the functioning of the RNR complex and thus might alter the balance of stoichiometries.

If the RNR complex exists in a range of stoichiometries as suggested by the purification data from Nestoras *et al.*, (2010) (Figure 4.7) it would be expected that a broad maximum entropy peak would be observed. This peak would cover all the possible values from a $Cdc22^{R1}$ monomer (+mEos3.2) upwards. The third peaks seen are not as broad as might be expected for this explanation, being similar in width to the peak presumed to represent cleaved mEos3.2. However, there is sufficient variability in the third peak it would seem unwise to rule out a spectrum of stoichiometries.

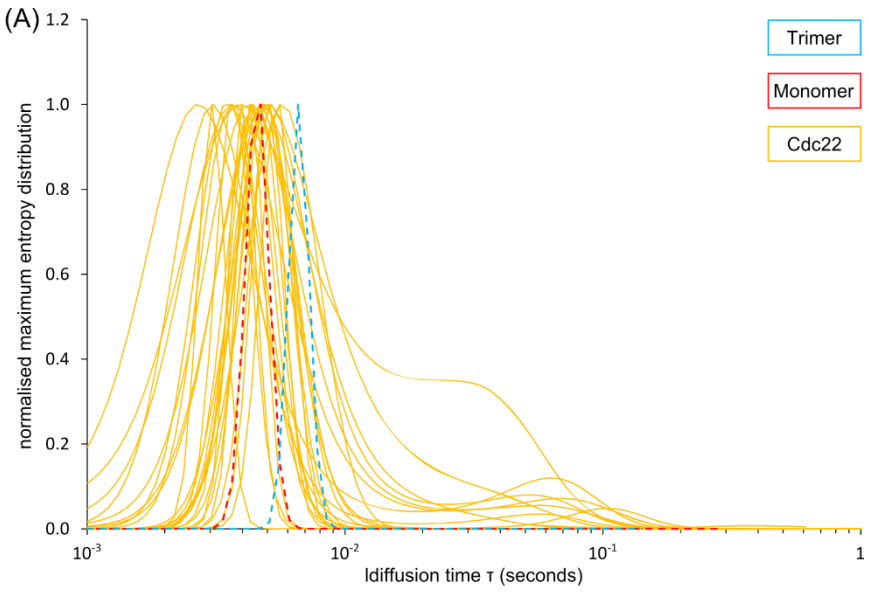
Calculating the diffusion coefficient for the peak values of the low third peaks gives a diffusion coefficient of 0.5-3 $\mu m^2/s$ which is a ~5 fold decrease in speed compared to what was calculated for monomeric mEos3.1. This would suggest a 125 fold increase in MW compared to the monomer, assuming a roughly spherical protein. Possibly this indicates a fraction of $Cdc22^{R1}$ or cleaved mEos3.2 that is aggregating. Alternatively it might be an indication that the larger $Cdc22^{R1}$ complex is not spherical, having a much greater hydrodynamic radius than a spherical protein of its MW.

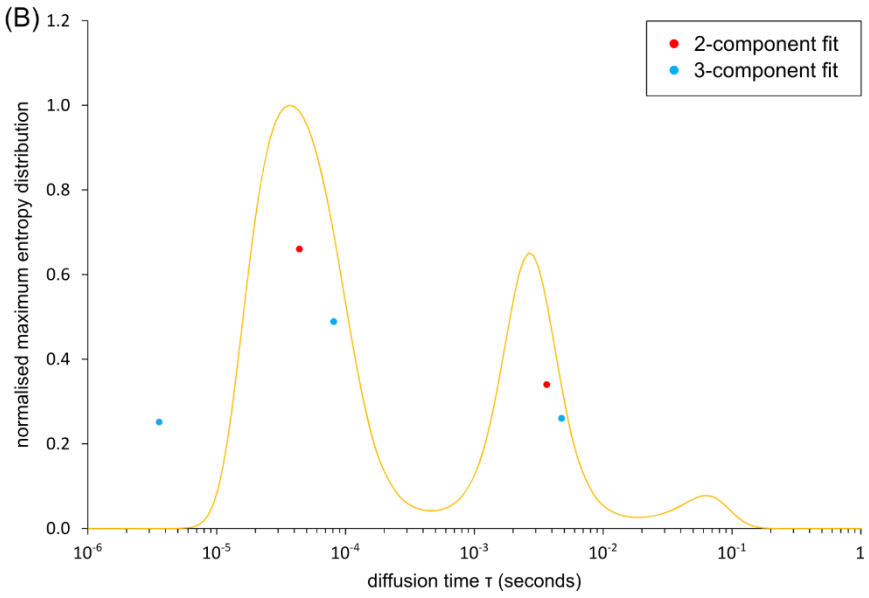
Figure 4.17 Examining the third peak in Cdc22^{R1} maximum entropy distributions

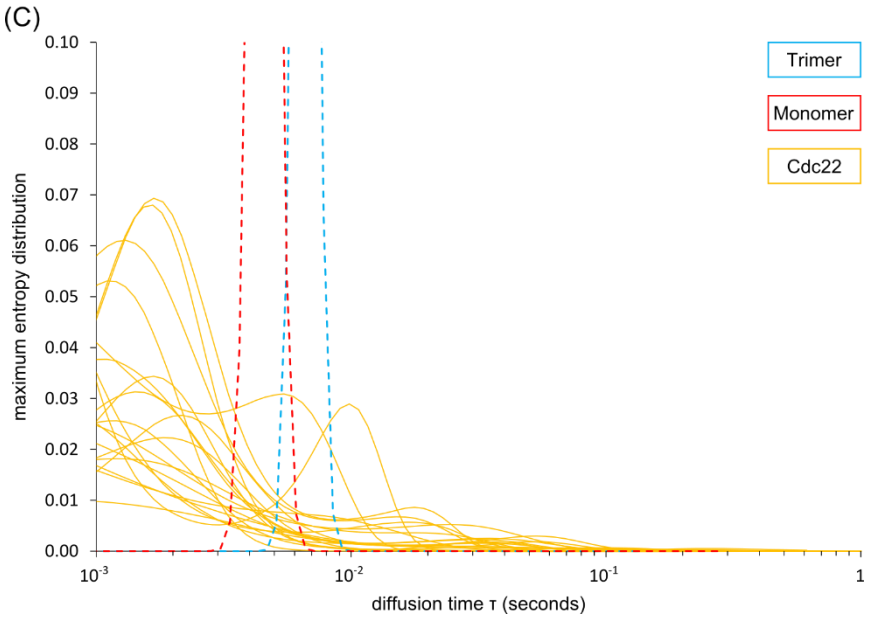
(A) The maximum entropy distributions for the Cdc22^{R1} data and a representative monomer and trimer distribution were all normalised such that the monomer-like peak was equal to 1. When this was done it revealed that a number of the distributions had a very low peak at a longer diffusion time. There is also a lot of variability within the Cdc22^{R1} data.

(B) A single distribution with three distinct peaks was plotted normalised to the tallest peak. The FCS curve that gave this distribution was then fitted using the conventional fitting method and either a two or a three component fitting model. The resulting diffusion times were plotted against their respective fractions over the maximum entropy distribution. It can be seen that three component fitting has not found the third peak in the maximum entropy distribution but has instead fitted a shorter diffusion time not corresponding to anything shown by the maximum entropy distribution.

(C) The Cdc22^{R1} FCS data were re-analysed to only include the section corresponding to the monomer-like peak and the third, long time peak in both the averaging and the fitting. This was achieved by only using the data from time 5×10^{-4} s onwards with the curve averaging algorithm. The re-analysis had the effect of shifting the monomer-like peak to the left, to a shorter time, and revealed a longer time peak for more of the distributions. However, there was a great variability in where the peaks fall making it hard to interpret this data.


(A)
Trimer
Monomer
Cdc22
normalised maximum entropy distribution
ldiffusion time τ (seconds)


(B)
2-component fit
3-component fit
normalised maximum entropy distribution
diffusion time τ (seconds)


(C)
Trimer
Monomer
Cdc22
maximum entropy distribution
diffusion time τ (seconds)

4.2.8.3 HU treatment affects the diffusion time of the RNR complex

One hypothesis is that the RNR complex varies its stoichiometry according to whether it is active or inactive, with the larger form (possibly $\alpha_6\beta_2$) being the inactive form, based on the size exclusion experiment by Nestoras *et al.*, (2010). To gain any information on the state of the complex from this method it was necessary to compare cells in which $Cdc22^{R1}$ would be behaving differently. FCS measurements were taken in cells that had been treated with 10 mM HU for 2.5 hrs (long enough for the majority of cells to be blocked in S phase). Measurements taken on two different days show reasonable consistency between them, and in fact more consistency than was seen between the untreated cells (Figure 4.18).

The right-most maximum entropy distribution peaks for these experiments had a longer diffusion time than either the monomer or the trimer (Figure 4.18). However, this was still the strain in which mEos3.2 is being cleaved from the majority of $Cdc22^{R1}$ so it is arguable whether this actually represents a complex. Another point worth noting is that (for the most part) the HU treated maximum entropy distributions did not show the third, long time peak seen in a number of the distributions for untreated cells.

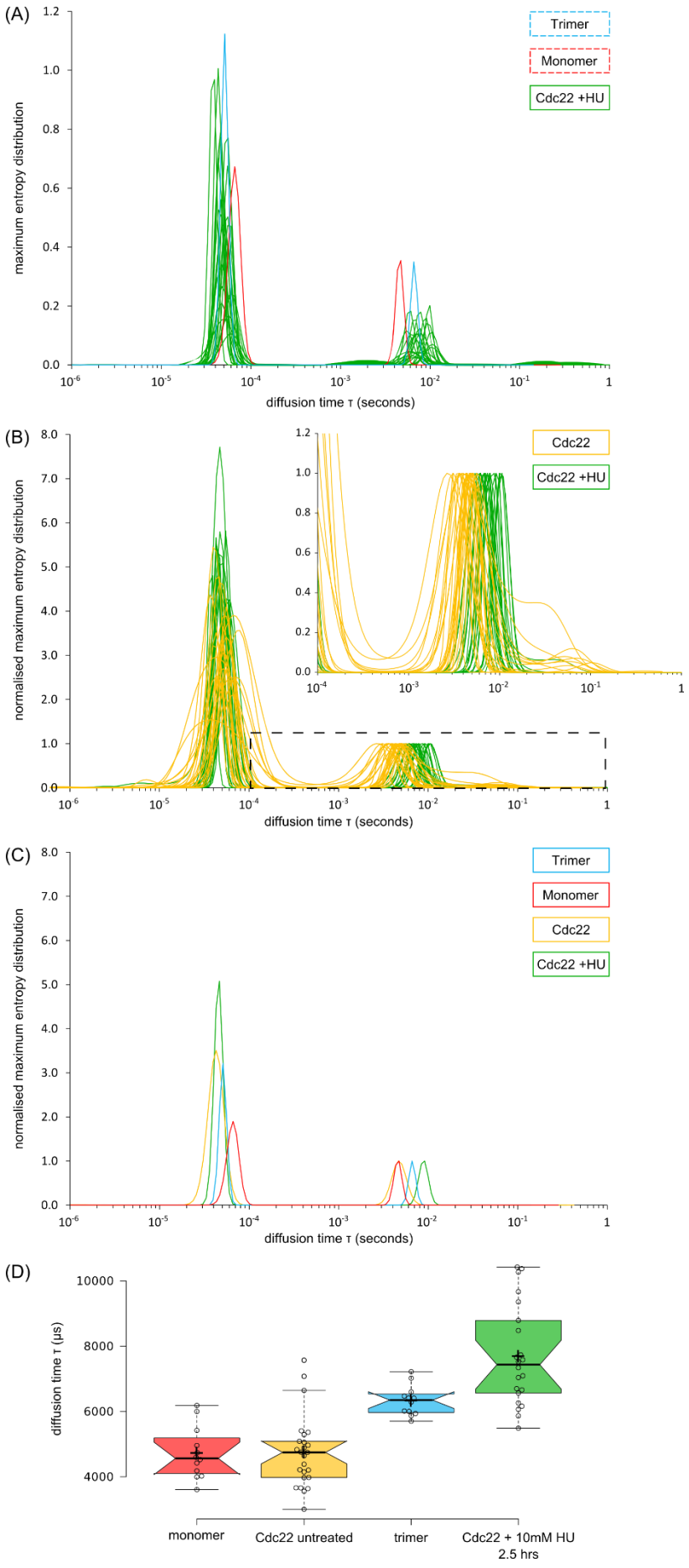
Figure 4.18 Treatment with HU changes the maximum entropy distribution for Cdc22^{R1}-mEos3.2

(A) The maximum entropy distributions for FCS curves obtained from cells treated with 10 mM HU for 2.5 hrs plotted with representative monomer and dimer distributions. As in all the mEos3.1 samples the distributions show a short time peak just below 1×10^{-4} s. For the HU treated cells the long time peak falls at a longer time than the trimer peak.

(B) The distributions were normalised to the second peak and plotted with similarly normalised distributions from untreated Cdc22^{R1} mEos3.2 cells. The majority of the HU treated distributions do not show a third peak when plotted this way whereas a number of the untreated distributions do.

(C) For increased clarity representative single distributions from treated and from untreated cells were plotted along with the monomer and trimer distributions, all normalised to the right-most peak. It can be seen that the peak for untreated Cdc22^{R1} mEos3.2 sits right on top of the peak for monomer and the peak for HU treated Cdc22^{R1} mEos3.2 sits at a longer time than the peak for trimer.

(D) A boxplot of the diffusion times from conventional fitting with a two component curve gives the same relationships.


(A)
Trimer
Monomer
Cdc22 +HU
maximum entropy distribution
diffusion time τ (seconds)
(B)
Cdc22
Cdc22 +HU
normalised maximum entropy distribution
diffusion time τ (seconds)
(C)
Trimer
Monomer
Cdc22
Cdc22 +HU
normalised maximum entropy distribution
diffusion time τ (seconds)
(D)
diffusion time τ (μs)
monomer
Cdc22 untreated
trimer
Cdc22 + 10mM HU
2.5 hrs

4.3 Discussion

4.3.1 The limits of measuring very slow diffusion

It is possible that only a very little of the slow third peak in $Cdc22^{R1}$ was observed for two reasons. Firstly, the FCS signal is very much dominated by the faster time component of the photo-physical effect. Secondly, the complex may be moving so slowly that a majority of it is bleaching before it can cross the detection volume. One possible way to test this would be to repeat the experiment with purified mEos3.1 in solution in increasing concentrations of glucose. If, as the molecule slows, a saturation in the diffusion time is reached that would give an indication of the limits of mEos3.1 for this purpose. In addition, purified mEos3.1 in a high concentration solution could be measured with everything in the detection volume activated such that the FCS measurement would track the changes from photo-bleaching rather than diffusion. If it were discovered that mEos3.1 bleaches faster than a large $Cdc22^{R1}$ complex is expected to be diffusing then it would be necessary to look into using a fluorophore with a longer photo-bleaching time. Unfortunately, the majority of fluorescent proteins photo-bleach quite rapidly. Organic fluorescent dyes might be an answer as they photo-bleach more slowly, but it can be harder to introduce organic dyes into cells (see Introduction 1.2.3.1).

4.3.2 Photon count histogram analysis of oligomerisation

In 1999 Chen *et al.*, demonstrated a novel analysis method for determining numbers of molecules in FCS called the photon count histogram (Chen *et al.*, 1999). They showed that photon count histogram analysis could be used to separate a mixture of fluorophores based on brightness (Müller *et al.*, 2000). They then went on to prove that this analysis technique is useful for studying oligomerisation in biological systems because the

brightness of a species scales linearly with the number of fluorophores it is tagged with (Chen *et al.*, 2003). Using enhanced green fluorescent protein (EGFP) they were able to examine the oligomerisation states of two receptor proteins; the testicular receptor 4 and the retinoid X receptor. Photon count histogram analysis has the potential to be very helpful in studying of the oligomeric states of $Cdc22^{R1}$.

However, it is not likely to be easily applied to mEos3.2 tagged $Cdc22^{R1}$ for a few reasons. Firstly, it would be necessary to activate all of the mEos3.2 molecules on a single complex at the same time in order to observe the photon count increase. To achieve this it would be necessary to use a higher power of activation 405 nm laser and run the risk that the average number of fluorescent molecules in the detection volume is also increased, swamping out the signal. Secondly, it has been demonstrated that mEos3.2 has a relatively poor photo-activation efficiency of ~40% (Durisic *et al.*, 2014) meaning that it may not even be possible to see fluorescence from every $Cdc22^{R1}$ subunit in the complex. Thirdly, evidence has been shown here that suggests a portion of the mEos3.2 fluorophore is being cleaved from the $Cdc22^{R1}$ (see 4.3.7.1) meaning that it is unlikely that every $Cdc22^{R1}$ subunit in a complex has an mEos3.2 tag. Should a more stable tag be found with a better activation efficiency this analysis technique might be worth pursuing, although the problem of having too many fluorescent molecules in the detection volume would remain.

4.3.3 The effect of HU on the RNR complex

One interesting point to note is that, in the untreated $Cdc22^{R1}$, FCS maximum entropy distributions sometimes show a peak at a longer diffusion - albeit a very low peak. In the data from cells treated with the RNR inhibitor HU there is only one maximum entropy distribution that shows a similar peak. Nestoras *et al.*, (2010) observed that cells treated with HU did not have a FRET signal between $Cdc22^{R1}$ and $Suc22^{R2}$. One possible reason

for this is that HU inhibits RNR complex formation and that the low third peak in the untreated cells does represent a form ($\alpha_2\beta_2$ or $\alpha_6\beta_2$) of the RNR complex. However, *in vitro*, HU does not disrupt RNR complexes and Nestoras *et al.*, also observed in size exclusion chromatography that HU treatment shifted the size of eluted product towards larger complexes. Thus HU treatment might be expected to result in more of the long time peak. Combining this with the fact that in the HU-treated cells a peak is not seen that overlaps with the monomer peak it may be that HU treatment has actually increased the quantity of the long time peak component with the effect that the two maximum entropy distribution peaks (the short time cleaved mEos3.2 peak and the long time putative complex peak) are merged into a single peak that sits between the two.

4.3.4 The effects of the cell cycle on the RNR complex

It has been shown that the transcript level of the R1 subunit is regulated in a cell cycle dependent manner, with the highest transcript level occurring around G1/S phase (Björklund *et al.*, 1990; Gordon and Fantes, 1986) (see Introduction 1.3.1 for a description of the *S. pombe* cell cycle). This makes sense: the dNTP pools would need to vary according to whether the cell was actively replicating its DNA or not. It is also likely that the activity of the RNR complex would be regulated in a cell cycle dependent manner for the same reason. If it is assumed that there exist two or more different stoichiometries of the RNR complex with different activity levels it might be expected that a difference in the proportions of each stoichiometry would be seen at different points in the cell cycle. As such it would be an interesting experiment to use a lactose gradient to enrich for small G2 cells and compare FCS measurements of $Cdc22^{R1}$ mEos3.2 in cells in the G2 and S phases of the cell cycle. Based on the cell cycle regulation of RNR activity it would be expected that more of the active complex would be present in S phase than in G2, which might lead

to a shift in peak position which could be observed using the maximum entropy fitting.

4.3.5 Other RNR regulatory factors

In *S. cerevisiae* one of the factors found to regulate RNR activity was a small, intrinsically disordered protein called Sml1. Sml1 was first identified by Zhao *et al.*, (1998) when they found a mutant strain which was able to grow even in the absence of the essential gene *MEC1*. It was shown that in Δ*SML*1 strains the dNTP pool levels were elevated and that this effect was not due to an increase in transcription of the RNR proteins. It was also found that Sml1 binds to the R1 subunit of the RNR complex (Zhao *et al.*, 1998). It was subsequently shown by Chabes *et al.*, (1999) that, *in vitro*, purified Sml1 inhibits RNR activity as measured by a CDP reduction assay (Chabes *et al.*, 1999). They also used a BIAcore binding experiment using R1 titrated against immobilised Sml1 to show that there is a binding interaction. They then used sucrose gradient centrifugation to test whether Sml1 binding would affect the stoichiometry of the R1 protein. As controls they examined R1 + dTTP which gave a mixture of dimeric and multimeric R1, and R1 alone which gave mainly monomeric R1. Adding Sml1 did not shift the distribution of oligomers seen.

In *S. pombe* a similar peptide called Spd1 was identified (Woollard *et al.*, 1996), this protein has been shown to bind to $Cdc22^{R1}$ and also inhibit RNR activity (Hakansson *et al.*, 2006; Liu *et al.*, 2003) and has also been shown to be necessary for FRET between $Cdc22^{R1}$ and $Suc22^{R2}$ (Nestoras *et al.*, 2010). As such, it could potentially have some role in the formation of the distinct RNR complexes from $Cdc22^{R1}$ and $Suc22^{R2}$. Nestoras *et al*, (2010) used a cross-linked *S. pombe* protein extract that was run through a gel filtration column to see whether two discretely sized populations could be seen. The peak they saw from the gel filtration fractions, visualised on a western blot, covered a broad range of possible complex sizes suggesting a continuum rather than discrete populations (Figure 4.19). This may

possibly be an artefact of the crosslinking or it may suggest that the complex actually exists in a continuum of oligomeric states rather than discrete populations. They also compared a wild-type extract to an extract from an *spd1* delete mutant and showed that the tail of the fraction peak was shifted towards the higher weights. However, in HU treated cells the presence or absence of Spd1 had no effect on the weight distribution seen.

In light of these results it has been hypothesised that Spd1 promotes the formation of the smaller active RNR complex and possibly inhibits the formation of an inactive hexameric complex of $Cdc22^{R1}$ and in the absence of Spd1 more of the hexameric $Cdc22^{R1}$ is formed which, when bound to a dimer of $Succ22^{R2}$, does not give a FRET signal. For the future it would be interesting to cross the $Cdc22^{R1}$ base strain with an *spd1* delete strain and then cassette exchange in the mEos3.2 tag, take FCS measurements and compare the maximum entropy and conventional fitting to what was seen in the *spd1* wild type strain.

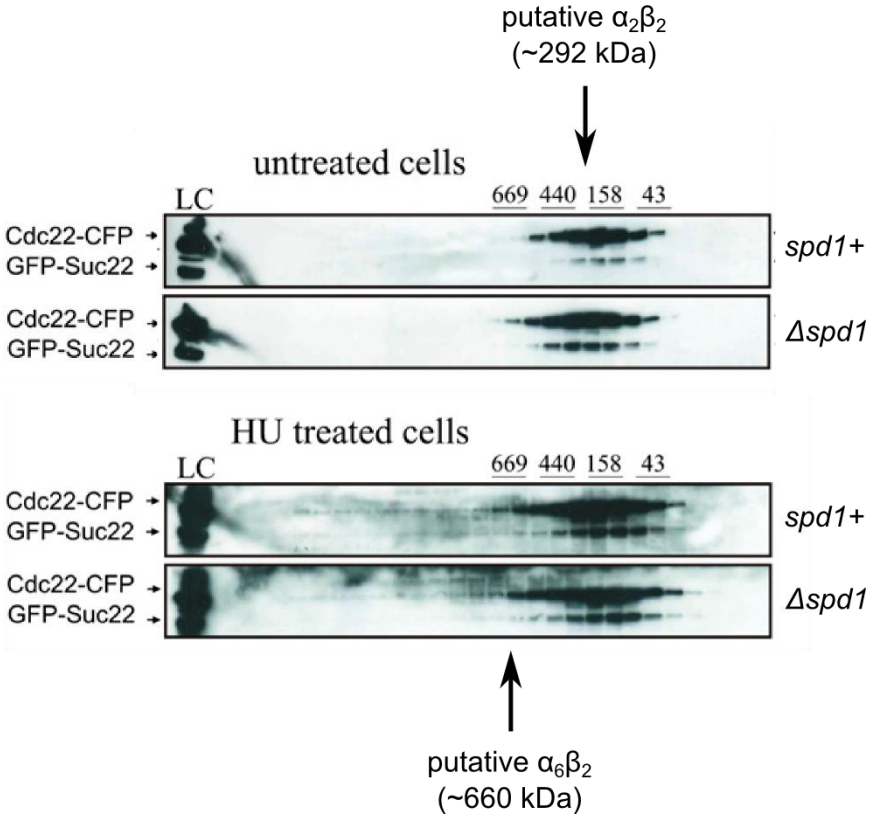


Figure 4.19 Size exclusion chromatography of HU treated RNR extract

Taken from Nestoras *et al.*, (2010)

Nestoras *et al.*, extracted RNR proteins from *S. pombe* cells with and without HU treatment. Using size exclusion and a western blot of the collected fractions Nestoras *et al.*, showed that in the Δ*spd1* (delete) cells there was a slight increase in the concentration of $Cdc22^{R1}$ in fractions at size suggesting the $\alpha_6\beta_2$ complex. In cells that had been treated with HU the presence or absence of *spd1* did not seem to have an effect on complex size.

4.3.6 The usefulness of the mEos family of fluorescent proteins for FCS

The photo-activatable properties of the mEos family of fluorescent proteins are very desirable for FCS, especially when working with proteins with a high expression level, as they give control over the number of fluorescent molecules in the detection volume. However, the mEos family of proteins seem to have a number of downsides as fluorescent tags. Firstly, it has been found that mEos3.2 only has a photo-activation efficiency of 40% (Durisic *et al.*, 2014) (mEos2 is slightly better at 60%). However, this does not matter as much when dealing with a protein that is as highly expressed as $Cdc22^{R1}$ because it is still possible to activate sufficient numbers of mEos3.2 molecules to obtain FCS data. Secondly, although mEos3.1 and 3.2 are among the brighter of PA-FPs, they are not as bright as the organic dyes available, thus the SNR is not optimal. Additionally they photo-bleach somewhat rapidly compared to organic dyes, possibly on a timescale that matters (see 4.3.1). Thirdly, although the mEos family of proteins are thought to be structurally similar to GFP, consisting of a chromophore protected by a beta barrel, other members of the Carr group have had difficulty tagging a protein with mEos3.1/3.2 where it has been previously successfully tagged with GFP (data unpublished).

The first reported Eos protein (Jörg Wiedenmann *et al.*, 2004) was not monomeric and would spontaneously tetramerise in solution. This was thought to be problematic for studying protein behaviour *in vivo* so various mutated forms were designed with the aim being to retain the fluorescent properties but render Eos monomeric (McKinney *et al.*, 2009; Mingshu Zhang *et al.*, 2012) mEos2 and mEos3.1 and 3.2 respectively. However mEos2 was found not to be truly monomeric and the mEos3s were found to have a lower photo-activation efficiency than mEos2. Tagging $Cdc22^{R1}$ with mEos3.2 had to be attempted by multiple techniques and even once it was achieved, the cells were found to

cleave the mEos3.2 tag from $Cdc22^{R1}$. This very strongly suggests the mEos3.2 tag was interfering with $Cdc22^{R1}$ activity.

Ideally for FCS, a fluorophore would be found that would be very bright, easy to tag $Cdc22^{R1}$ with in a 1:1 ratio without affecting function and which would photo-bleach very slowly.

4.4 Summary

A combination of organic fluorescent dyes and purified mEos3.1 have been used to characterise an FCS system with a view to applying FCS to measure RNR complex stoichiometry. The maximum entropy fitting algorithm has been compared to the canonical fitting method. It has been established that using FCS and a combination of conventional fitting and maximum entropy fitting it is possible to distinguish between a monomer, dimer and trimer of mEos3.1 *in vivo*. Using different sized rhodaminated dextrans it was shown that it is harder to distinguish two different populations in a mixture using either fitting method. Looking at $Cdc22^{R1}$ tagged N-terminally with mEos3.2 using western blot analysis it has been found that the majority of the tag is cleaved. However, there did appear to be a small population of mEos3.2 molecules diffusing at a very slow speed, possibly indicative of the large RNR complex. By treating the $Cdc22^{R1}$ N-terminal mEos3.2 cells with HU it was hoped that a difference in maximum entropy profile would be seen. It was observed that the peak at the longest time was absent in the majority of HU treated cells and that the peak for cleaved mEos3.2 appeared to be shifted. From this data it is difficult to draw any conclusions about the size of the RNR complex. However it does suggest that FCS is a potentially useful technique for studying complex stoichiometry should a better fluorescent tag for $Cdc22^{R1}$ be found.

5. Using PALM to study the low abundance protein Rrp2 at endogenous levels

5.1 Introduction

Another potential use for photo-activated localisation microscopy (PALM) is the study of low abundance proteins that cannot be easily seen or analysed using conventional fluorescence microscopy. The DNA damage repair protein Rrp2 has been chosen as an appropriate example to explore the utility of PALM-based approaches for low abundance proteins. Rrp2 has been characterised as a gene involved in Rad51-dependent homologous recombination (HR). Thus, in this chapter the mechanism of Rad51-dependent DNA damage repair is introduced followed by a description of Rrp1 and Rrp2. The results of PALM imaging of fixed cells expressing Rrp2-mEos3.1 with and without DNA damage are presented. This is followed by the use of motion-blur PALM to provide a quantification of the fraction of Rrp2 associating with the DNA with and without DNA damage in a variety of different backgrounds. The results are discussed in light of unpublished data from two other groups and further experiments are suggested.

5.1.1 DNA damage and homologous recombination repair

5.1.1.1 Replication fork stalling and collapse

In eukaryotes DNA replication begins at so-called replication origins. When an origin fires (replication starts) structures called replication forks are formed which progress in each direction away from the origin. The replication fork consists of two arms of unwound ssDNA that act as the replication template and a single arm of dsDNA that is yet to be unwound. In addition, the replication complex, or replisome, contains the various proteins

needed for replication to occur, these include a helicase, a primase, leading and lagging strand polymerases, PCNA and a variety of other enzymes and structural proteins. The replication fork progresses as the helicase unwinds the DNA. There are a number of potential blocks to replication fork progress (see Introduction: 1.3.3.1); leading to a stalled fork in the short term and, potentially, a collapsed fork in the long term. Collapsed forks are forks which are unable to resume replication and can lead to DNA damage including double strand breaks (DSBs). Cells can cope with replication fork stalling either by stabilising the fork such that it can resume replication when the block is cleared or else by recovering (re-starting) collapsed forks. Both processes have been linked to HR proteins.

The fork stabilising process has been shown to depend on components of the S phase checkpoint (reviewed in Branzei and Foiani, 2010) (see Introduction: 1.3.3.3 for a description of checkpoint pathway activation). At stalled replication forks the activation of Rad3 initiates the S phase checkpoint signalling cascade, which inhibits cell cycle progression and prevents further origin firing. This helps to prevent damaged or un-replicated DNA being passed on to daughter cells and preserves the resources required to finish DNA synthesis. The S phase checkpoint also prevents the helicase unwinding at already active forks from becoming uncoupled from replication and thus creating too much unstable ssDNA. Arrested fork stabilisation has been shown to be dependent on recombination proteins (Costanzo, 2011) and replication re-start has been shown to occur by recombination-dependent pathways (Lambert *et al.*, 2010)

In both *E. coli* (Michel *et al.*, 1997) and mammalian (Arnaudeau *et al.*, 2000) cells inhibition of replication has been shown to cause DSBs and induce HR. In 2002 Lundin *et al.*, showed that replication forks arrested or slowed with HU treatment form DSBs at newly synthesised DNA and that HU treatment increased the frequency of recombination (Lundin *et al.*, 2002). The same group showed that overexpression of the HR protein

Rad51 in *S. cerevisiae* lead to a reduction of etoposide induced DSBs (Lundin *et al.*, 2003). In mouse cells replication inhibition leads to accumulation of DSBs and formation of Rad51 foci (Saintigny *et al.*, 2001). Petermann *et al.*, (2010) showed that, in mammalian cells, a long treatment with HU, the ribonucleotide reductase inhibitor that stalls DNA replication by depleting the dNTP pool, led to an accumulation of DNA DSBs (Petermann *et al.*, 2010). The same long treatment with HU also led to an accumulation of Rad51 in foci, which is indicative of HR. However, Petermann *et al*, 2010 also showed that at shorter treatment times DSBs did not occur, although recombination proteins were still involved in cell survival. In mammalian cells camptothecin can also induce DSBs at replication forks and mutants deficient in HR have been shown to be more sensitive to this type of damage (Arnaudeau *et al.*, 2001). All of these data suggest that HR is involved in the recovery of collapsed forks, in many cases linked to the repair of DSBs.

5.1.1.2 Double strand breaks

When a DSB in DNA is formed it must be repaired before replication can continue. There are two major pathways for this, HR and non-homologous end joining (NHEJ). In HR nuclease resection of the 5' ends at the break gives rise to a 3' single strand DNA (ssDNA) tail. This structure on its own is unstable but can be stabilised by the binding of replication protein A (RPA), which binds with a high affinity (Figure 1.11). In order for DSB repair to continue, and for recombination to occur, Rad51 needs to access the ssDNA tail where it will bind and form a nucleo-protein filament. The Rad51 nucleoprotein filament is key to promoting strand exchange. Rad51 assembles on ssDNA with the help of various mediator proteins and complexes. Rad52 is needed to target Rad51 to complexes of RPA (New *et al.,* 1998) and can be seen to form foci (clusters of multiple molecules) on the DNA in response to damage. The Rad55/Rad57 heterodimer complex (S. *cerevisiae RAD55/57* are

directly conserved in *S. pombe* and have othologs in mammalian cells known as the "Rad51 paralogs") is needed to stabilise the Rad51 nucleo-protein filament in order to enhance strand exchange (Sung, 1997).

In *S. pombe* the Swi5/Sfr1 complex has the same function as Rad55/57 and acts in parallel with but independently of it (Akamatsu *et al.,* 2007). Once loaded Rad51 promotes a search for homologous sequences and promotes strand exchange to ensure correct repair of the double strand break. The *rad51* gene in *S. pombe* was first identified by heterologous hybridisation using homology to the *S. cerevisiae RAD51* gene (Muris *et al.*, 1993). Mutants in which *rad51* has been deleted are very sensitive to damage by ionising radiation (IR), highlighting the importance of HR in DSB repair in *S. pombe*. The Rad51 null mutants are also sensitive to the chemicals hydroxyurea (HU) and methyl methanesulfonate (MMS). HU is an inhibitor of the ribonucleotide reductase complex (Lammers and Follmann, 1984), which is required for the production of deoxy nucleotide triphosphates, the building blocks of DNA. By inhibiting the complex HU depletes the pool of dNTPs and stalls DNA replication in S phase. When replication is resumed a small number of the replication forks will collapse and HR is required for the recovery of these collapsed forks. MMS is an alkylating agent which causes DNA damage mainly in S phase, it is assumed this damage is a result of a collapsing fork (Lundin *et al.*, 2005) although the exact mechanism remains unknown.

5.1.2 The Rad55/Rad57 pathway

The *S. pombe* HR repair gene *rad55* was identified through homology to the bacterial recA gene (homologous to *RAD51*) but was found to have most homology to *S. cerevisiae RAD55* so was initially named *rhp55* (rad homolog in *pombe*) (Khasanov *et al.,* 1999). The *rad57* gene (initially *rhp57*) was identified by a screen for mutants with hypersensitivity to

MMS and IR, and synthetic lethality with a *rad2* mutation. It was found to be epistatic with *rad51* (Tsutsui *et al.*, 2000). Essential for crossover production (Akamatsu *et al.,* 2007), the Rad55 and Rad57 proteins form a stable heterodimer (Sung, 1997) and increase the strand exchange efficiency of Rad51. Whilst in *S. cerevisiae RAD55* and *RAD57* null mutants are as sensitive to ionising radiation (IR) as *RAD51* null mutants in *S. pombe* the *rad55* and *rad57* null mutants are only partially sensitive when compared to *rad51* null mutants.

5.1.3 The Swi5/Sfr1 pathway

Fission yeast *swi5* was first described in 1984 when it was found as part of a screen for *S. pombe* mutants deficient in mating type switching (Stern *et al.,* 1984). The *swi5* gene was subsequently shown to be conserved in eukaryotes and the Swi5 protein in *S. pombe* was found to be involved in a Rad51-dependent recombination repair pathway that was independent of Rad55/57 (Akamatsu *et al.,* 2003). Swi5 was also shown to have two distinct roles; interacting with Sfr1 in its recombination repair function and Swi2 in its mating type switching function (Akamatsu *et al.,* 2003). The *swi2* gene was initially identified through its requirement for mating type switching in *S. pombe* and the *sfr1* gene was discovered by a BLAST search for sequence similarity to *swi2* (Akamatsu *et al.*, 2003). Swi5 and Sfr1 have been shown to localise to the nucleus in mitotic cells and to form foci in the nucleus in response to damage (Akamatsu *et al.,* 2007). The Swi5-Sfr1 complex has also been shown to stimulate Rad51-mediated strand exchange *in vitro* (Haruta *et al.*, 2006).

Mutants in which *swi5* and *sfr1* have been deleted are mildly sensitive to IR but double mutants with *rad55* or *rad57* deleted as well are as sensitive as Δ*rad51*. This indicates that Swi5/Sfr1 and Rad55/57 act in parallel pathways to mediate Rad51-dependent strand exchange.

5.1.4 Discovery of *rrp1* and *rrp2*

The *S. pombe* DNA repair proteins Rrp1 and Rrp2 were first identified by Dziadkowiec *et al.,* (2009) who were looking for sequences in *S. pombe* with homology to the C-terminal end of the *S. cerevisiae* protein Ris1, which is a DNA-dependent ATPase with roles in DNA repair and gene silencing. The genes discovered showed 34% (*rrp1*) and 36% (*rrp2*) sequence similarity to the C-terminus of Ris1 and 41% similarity to each other. These proteins are paralogs of each other and had been predicted to have helicase and DNA-dependent ATPase activity. Both proteins contain a zinc finger C3HC4 type RING finger (not found in Ris1), a Helicase-C domain and an SNF2-N domain (both also found in Ris1), possibly suggesting they are members of the SNF2 family. Both genes also showed high similarity with *S. pombe rad8*, a member of the SNF2 family that encodes an ATPase with an ubiquitin protein ligase E3 function.

5.1.4.1 Role of *rrp1* and 2 in the Swi5/Sfr1 damage repair pathway

In spot tests deletion mutants of *rrp1* and *rrp2* (Δ*rrp1* and Δ*rrp2*) did not show any reduction in viability or sensitivity to treatment when tested with the chemical HU (or other chemicals which can induce DNA damage) in comparison to a wild type (Dziadkowiec *et al.,* 2009). Nor did a double mutant of Δ*rrp1* Δ*rrp2*. Due to the amino acid sequence homology with Ris1 it was presumed that Rrp1 and Rrp2 were involved in the Rad51-dependent DNA damage repair pathway. To establish which arm of the Rad51 pathway Rrp1 and Rrp2 might function in Δ*sfr1* and Δ*rad57* mutants were combined with Δ*rrp1* and Δ*rrp2*. Increased DNA damage sensitivity was observed in the Δ*rad57* Δ*rrp1*/Δ*rrp2* double deletion mutants and not in the Δ*sfr1* Δ*rrp1*/Δ*rrp2* double deletion mutants, suggesting that Rrp1 and Rrp2 were part of the Swi5/Sfr1 pathway rather than

the Rad55/57 pathway.

5.1.4.2 Formation of nuclear foci in response to damage

As Swi5 and Sfr1 had been shown to form foci in the nucleus in response to DNA damage it was hypothesised that Rrp1 and Rrp2 would also form foci in the nucleus in response to DNA damage. After treatment for 1 hour with 0.1% MMS, Dziadkowiec *et al.,* observed the formation of nuclear foci in strains overexpressing GFP or RFP tagged Rrp1/2 from a plasmid (Figure 5.1) (Dziadkowiec *et al.,* 2009, Dziadkowiec *et al.,* 2013). Rrp1 and Rrp2 were found to co-localise in these foci and also to co-localise with Rad52. The formation of these foci was not dependent on the presence of Rad51, Rad57 or Swi5. This focus formation and co-localisation suggests that Rrp1 and Rrp2 are recruited to sites of DNA damage and directly bind the chromatin. However, Dziadkowiec *et al.,* only observed this foci formation with Rrp1/2 overexpressed from a plasmid and were not able to detect the proteins at endogenous levels. It was thus hypothesised that, by using super-resolution microscopy techniques, it might be possible to improve the resolution and/or detection sufficiently to characterise these proteins at endogenous levels.

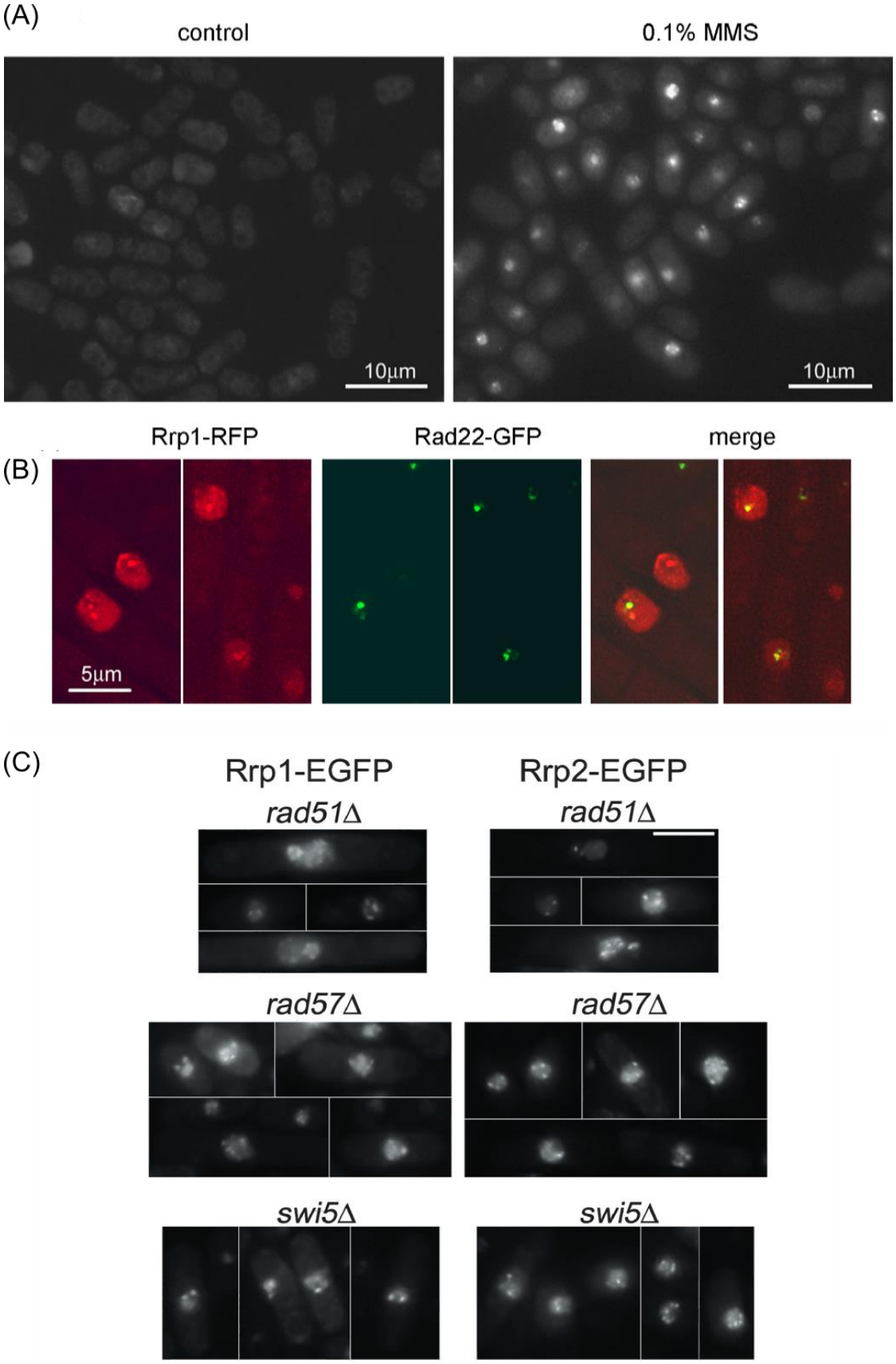
Figure 5.1 Foci formation of Rrp1 and Rrp2 after treatment with MMS

Taken from (Dziadkowiec *et al.,* 2009) (A & B) and (Dziadkowiec *et al.,* 2013) (C)

(A) RFP-Rrp1 expressed from a plasmid forms foci after treatment with 0.1% MMS for 1 hour (Dziadkowiec *et al*., 2009).

(B) Co-localisation of Rad52 with RFP-Rrp1 in MMS treated cells (Dziadkowiec *et al*., 2009).

(C) Formation of Rrp2 as well as Rrp1 foci does not require Rad51, Rad57 or Swi5 (Dziadkowiec *et al*., 2013).


(A)
control
0.1% MMS
10μm
10μm
(B)
Rrp1-RFP
Rad22-GFP
merge
5μm
(C)
Rrp1-EGFP
Rrp2-EGFP
rad51Δ
rad51Δ
rad57Δ
rad57Δ
swi5Δ
swi5Δ

5.1.5 Motion blur PALM

In 2014 the Carr lab published a description of a PALM technique which allowed visual separation of static and freely diffusing molecules (Etheridge *et al.*, 2014). The technique makes use of the fact that, at long exposure times, freely diffusing molecules will appear as a blur on the camera and not be localised by single molecule localisation software. Whereas, relatively static molecules, such as those associated with the chromatin, should appear as a localisable point spread function (PSF) spot (Figure 5.2). The methodology makes the assumption that static proteins seen within the nucleus would most likely be associated with the DNA or chromatin. As such this technique can be used to show differences in the population of a protein associating with the DNA or chromatin in response to different treatments, for example when visualising proteins localising to the nucleus in response to DNA damage.

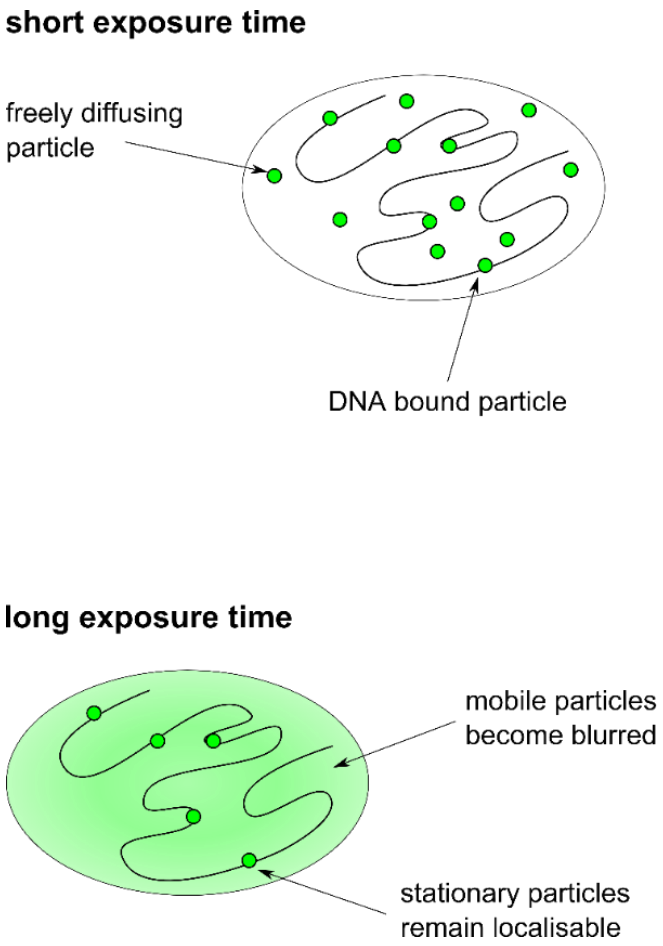


Figure 5.2 Motion Blur PALM

Adapted from (Etheridge *et al.*, 2014)

Particles in the nucleus that are associated with the DNA are likely to appear more "stationary" than freely diffusing particles. Freely diffusing particles can be viewed as essentially immobile spots at short exposure times, as in single particle tracking, but at long exposure times they will appear as a blur, rather than a spot, on the camera. In contrast stationary molecules will appear as a spot even at very long exposure times. Molecules that are DNA or chromatin associated will behave more like stationary molecules when exposure times are in the region of 200-500 ns (Etheridge *et al.*, 2014). This blurring effectively gives a tool to filter out particles that are not bound to the DNA: by imaging with a relatively long exposure time it is possible to localise only the associated particles. The number of localisations can then be quantified and compared across different treatments or strains for a relative measure of changes in DNA-association.

5.1.6 Objective: Can a very low abundance protein be studied at endogenous levels using PALM?

The objective of this work was to apply PALM techniques to a biological problem. The Rrp1 and Rrp2 proteins had so far only been visualised by widefield microscopy when over-expressed due to their very low abundance. It was hypothesised that the improved detection and resolution of PALM would allow characterisiation of even the very small number of molecules of Rrp1/2 that occur at endogenous levels. This is important because any observations made *in vivo* and at endogenous levels are more likely to be truly representative of the role of these proteins. This problem is more or less the opposite extreme of the study of the $Cdc22^{R1}$ subunit of the ribonucleotide reductase complex, which is highly abundant throughout the cytoplasm.

Within the Carr research group a PALM technique called "motion blur" had been developed to give a relative measure of the DNA or chromatin association of nuclear proteins under different conditions. This technique makes use of the fact that at long camera exposure times highly mobile proteins will give rise to a blurred spot that cannot be localised by single molecule software, whereas chromatin or DNA bound protein will appear as essentially static, giving a sharper spot that can be localised by single molecule software. Since the Dziadkowiec group had observed Rrp1 foci after damage it seemed likely that, with this motion blurr technique, more localisations would be seen (correlating with more protein associating with the DNA) after the induction of DNA damage.

5.2 Results

5.2.1 Creating the mEos3.1 tagged strains

5.2.1.1 Tagging Rrp2 through cassette exchange

As Rrp2 tagged with GFP at the C-terminus was functional (Dziadkowiec *et al.*, 2013) a C-terminal mEos3.1 strain was created. The *rrp2* gene replacement base strain was created as described in Watson *et al.,* (2008) by replacing the endogenous *rrp2* gene with a *ura4* marker gene flanked by the loxM and loxP sites needed for recombination (Figure 5.3 A) (see Appendix for table of primers used). To create a tagged construct for replacing the *ura4* gene in the base strain, the *rrp2* gene was first amplified from wild type genomic DNA with the stop codon removed and restriction sites for *Sph*I (at the N-terminal end), *Nhe*I, *Xma*I and *Sal*I (at the C-terminal end) added. The fragment was then cloned into the pAW8 plasmid, which contained the loxM and loxP sites, using *Sph*I and *Sal*I (Figure 5.3 B). The mEos3.1 gene was next amplified from a stock plasmid using primers to add the restriction sites *XmaI* and *Nhe*I and sub-cloned at the C-terminus of *rrp2* in the pAW8 plasmid (Figure 5.3 C). Sequencing was performed to check the gene and tag were both correct and in the same reading frame.

The plasmid was transformed (see 2.2.2) into the *S. pombe* base strain and transformants were selected for on EMM agar plates with no amino acid supplements. Cassette exchange was allowed to occur during overnight culture in liquid yeast extract media (Figure 5.3 D). Subsequently, colonies where cells had lost the *ura4* marker were selected for on 5-fluoroorotic acid (5-FOA) and checked with PCR for integration of *rrp2*-mEos3.1. A western blot was performed to check for protein expression. The expression levels of Rrp2 were so low that, using an antibody to the mEos family of proteins, tagged Rrp2 (~144 kDa) was not detectable (Figure 5.4). However it was possible to see mEos3.1 activation on

the PALM microscope (Figure 5.4).

5.2.1.2 Tagging Rrp1 through cassette exchange

The *rrp1-mEos3.1* strain was not as straightforward to create and required a slightly different approach. Initial attempts at creating the base strain for *rrp1* failed repeatedly so it was necessary to increase the lengths of both homologous arms. This involved re-designing plasmids to perform a fusion PCR with three segments; an "up" and a "down" which were amplified from genomic DNA extracted from AW310 and would form the homologous arms, and a "marker" segment (see Appendix for table of primers used). The marker segment was amplified from pAW1 and contained the lox sites and the *ura4* marker gene (Figure 5.5).

The Dziadkowiec group had found that Rrp1 was functional when tagged N-terminally, so the *rrp1* PCR fragment for cloning into pAW8 was created with a restriction site (*Nhe*I) for sub-cloning at the N-terminal (Figure 5.5) then cloned into the cassette exchange plasmid pAW8. The mEos3.1 sequence was then sub-cloned between the promoter and the start of the *rrp1* gene. The pAW8 *rrp1* N-terminal mEos3.1 plasmid was transformed into the *rrp1* base strain and transformants were selected for on EMM agar plates with no amino acid supplements. Cassette exchange was then performed in 10 ml YE cultures overnight. Subsequently colonies where cells had lost the *ura4* marker were selected for on 5-FOA and checked with PCR for integration of *rrp1-mEos3.1*.

Figure 5.3 Strategy for tagging *rrp2* with mEos3.1 at the C-terminal using Cre-lox recombination

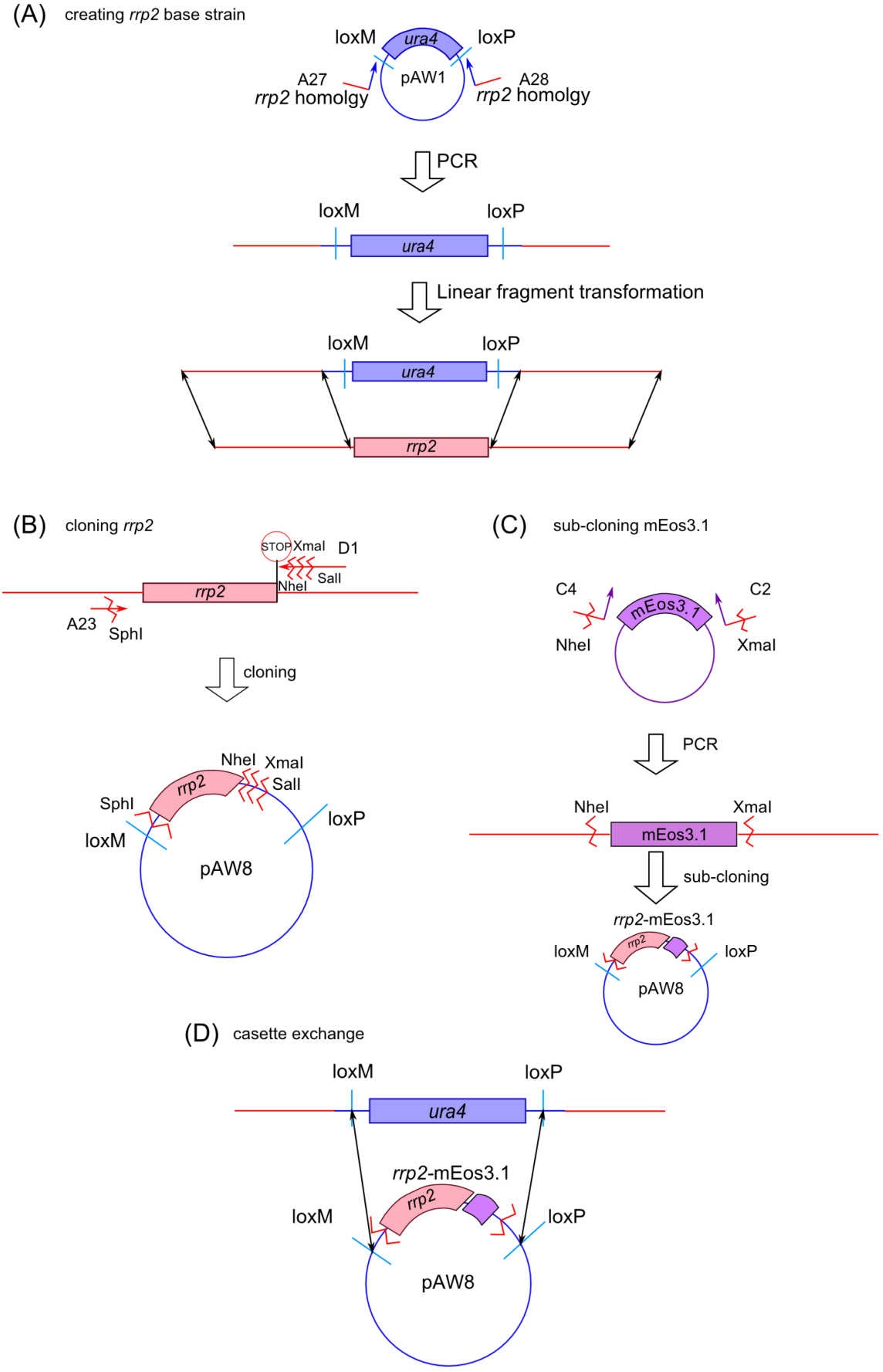
(A) Schematic for base strain construction by replacing the *rrp2* gene with the *ura4* marker gene flanked by loxM and loxP sites. Primers A27 and A28 (see Appendix) were used to amplify the *ura4* gene with 80bp arms homologous to the flanking regions of *rrp2*. The resulting fragment was transformed into a wild type strain where homologous recombination between the construct arms and the genome gives the base strain.

(B-D) Schematic for creation of rrp2 C-term mEos3.1 strain

(B) The *rrp2* gene was amplified from genomic DNA with the stop codon removed using primers A23 and D1. Restriction sites, *Sal*I, *Nhe*I, *Xma*I and *Sph*I, were added.

(C) The mEos3.1 gene was amplified from a plasmid with *Nhe*I and *Xma*I restriction sites using primers C2 and C4 then sub-cloned into pAW8 at the C-terminal end of *rrp2*.

(D) The pAW8 plasmid containing the mEos3.1 tagged *rrp2* gene was transformed into the base strain and cassette exchange was allowed to occur.


(A) creating rrp2 base strain
loxM
ura4
loxP
A27
rrp2 homolgy
pAW1
A28
rrp2 homolgy
PCR
loxM
loxP
ura4
Linear fragment transformation
loxM
loxP
ura4
rrp2
(B) cloning rrp2
STOP
XmaI
D1
NheI
SalI
rrp2
A23
SphI
cloning
NheI
XmaI
SalI
SphI
rrp2
loxP
loxM
pAW8
(C) sub-cloning mEos3.1
C4
C2
mEos3.1
NheI
XmaI
PCR
NheI
XmaI
mEos3.1
sub-cloning
rrp2-mEos3.1
rrp2
loxM
loxP
pAW8
(D) casette exchange
loxM
loxP
ura4
rrp2-mEos3.1
rrp2
loxM
loxP
pAW8

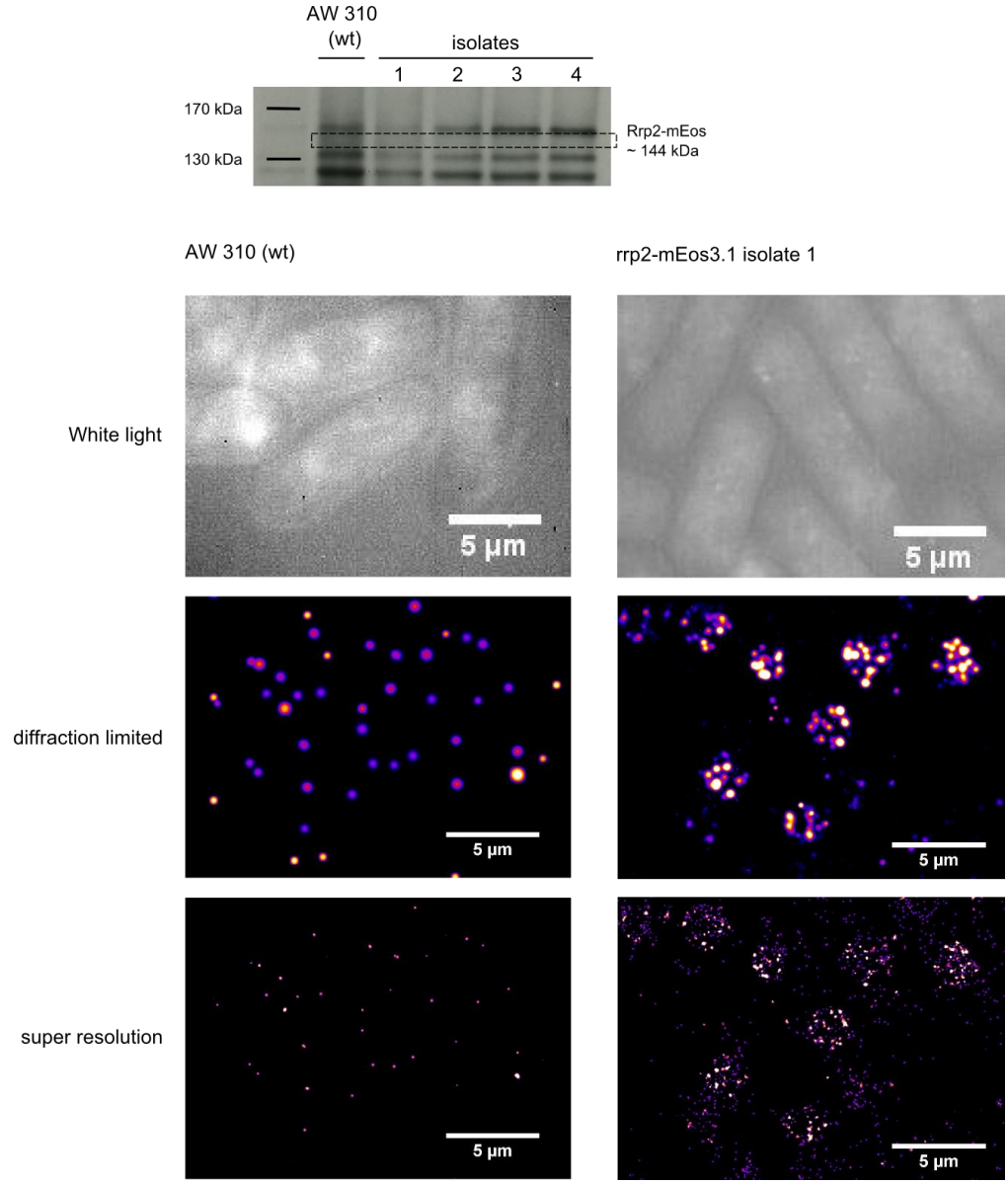


Figure 5.4 Visualising mEos3.1 tagged Rrp2

A western blot of various independent cassette exchange isolates compared to a wild type extract, treated with an antibody to the mEos family of proteins. The expected band should be at ~144 kDa, no band is detectable at this size.

Using the PALM imaging system it is possible to see photo-activatable fluorescence in the nuclei of cells. Localisations from *rrp2-mEos3.1* isolate 1 and wild type cells were plotted as a diffraction limited image and a super-resolution image where the spot width = the precision. There is a difference both in the number of localisations plotted and in the distribution of localisations.

Figure 5.5 Strategy for tagging *rrp1* with mEos3.1 at the N-terminal using Cre-lox recombination

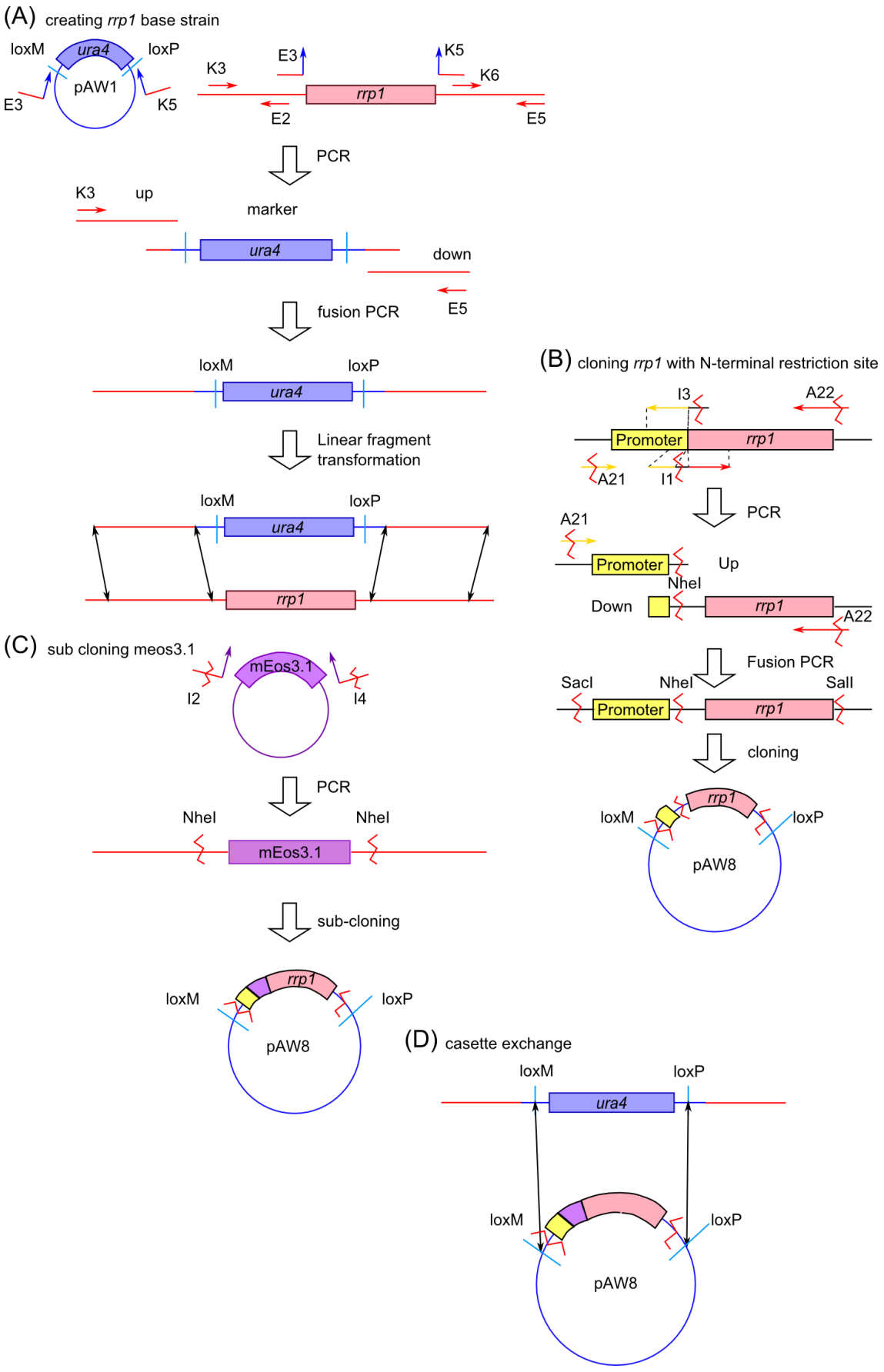
(A) Schematic for base strain construction by replacing the *rrp1* gene with the *ura4* marker gene flanked by loxM and loxP sites. Primers E3 and K5 (see Appendix) were used to amplify the *ura4* marker gene and lox sites, primers K3 and E2 and K6 and E5 were used to amplify long arms of homology which were combined with the marker fragment by fusion PCR. The resulting linear fragment was transformed into a wild type strain to give the *rrp1* base strain.

(B-D) Schematic for creation of *rrp1* N-term mEos3.1 strain

(B) The *rrp1* gene was amplified from genomic DNA in two fragments. One contained the promoter sequence with an *Nhe*I restriction site at the C-terminal end and a *SacI* site at the N-terminal end (primers A21 and I3). The other contained the *rrp1* gene with the *Nhe*I restriction site at the N-terminal end and a *Sal*I site at the C-terminal end (primers A22 and I1). The two fragments were combined by fusion PCR (using primers A22 and A23) and cloned into pAW8.

(C) The mEos3.1 gene was amplified from a plasmid with *Nhe*I restriction sites using primers I2 and I4 then sub-cloned into pAW8 between the promoter and the N-terminal end of *rrp1*.

(D) The pAW8 plasmid containing the mEos3.1 tagged *rrp1* gene was transformed into the base strain and cassette exchange was allowed to occur.


(A) creating rrp1 base strain
loxM
ura4
loxP
pAW1
E3
K5
K3
E3
K5
K6
rrp1
E2
E5
PCR
K3
up
marker
ura4
down
E5
fusion PCR
loxM
loxP
ura4
Linear fragment transformation
loxM
loxP
ura4
rrp1
(B) cloning rrp1 with N-terminal restriction site
I3
A22
Promoter
rrp1
A21
I1
PCR
A21
Promoter
Up
NheI
Down
rrp1
A22
Fusion PCR
SacI
NheI
SalI
Promoter
rrp1
cloning
rrp1
loxM
loxP
pAW8
(C) sub cloning meos3.1
mEos3.1
I2
I4
PCR
NheI
NheI
mEos3.1
sub-cloning
rrp1
loxM
loxP
pAW8
(D) casette exchange
loxM
loxP
ura4
loxM
loxP
pAW8

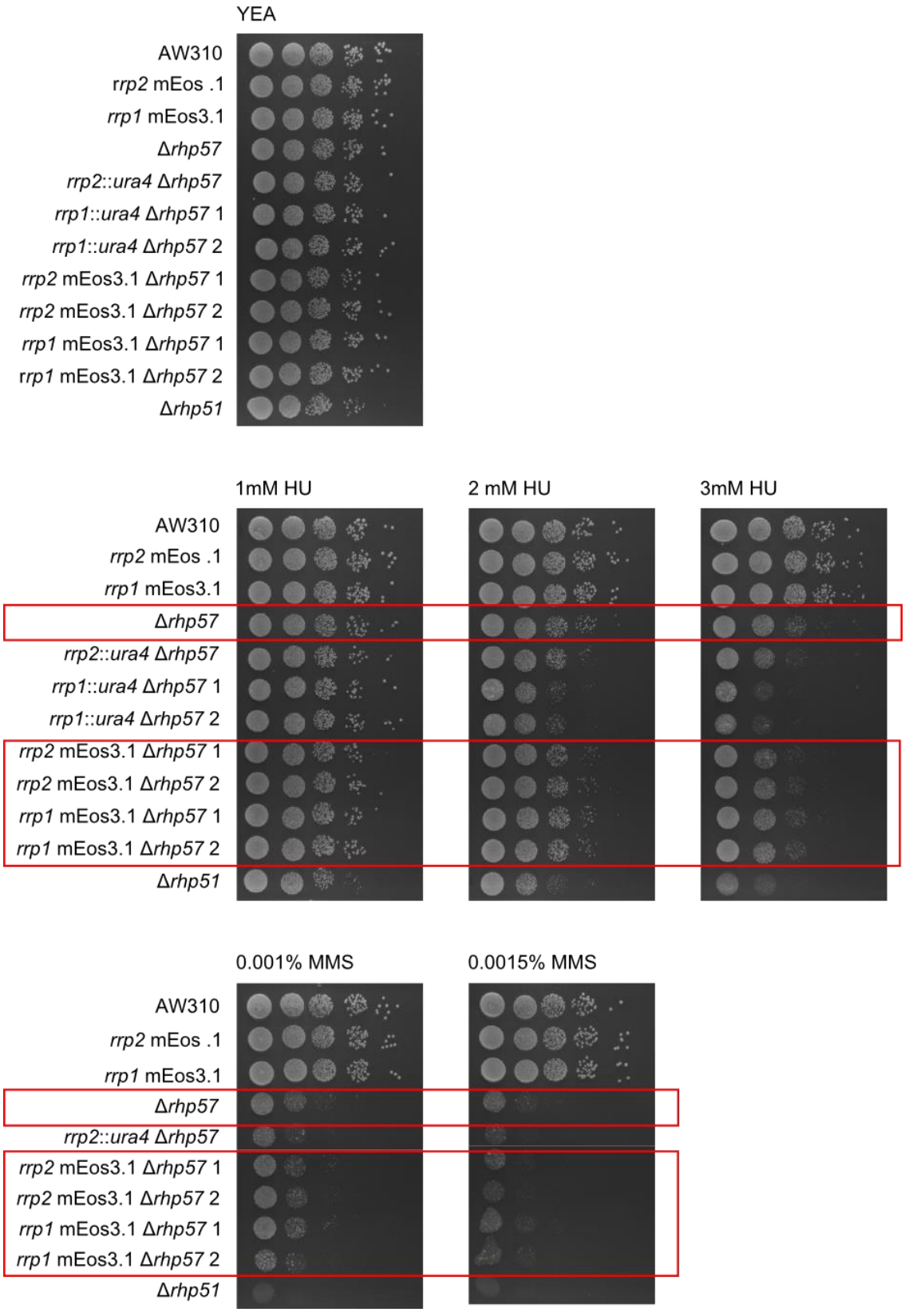
5.2.2 mEos3.1 tagged Rrp1 and Rrp2 are functional

Having successfully tagged both Rrp1 and Rrp2 with mEos3.1 it was necessary to check that they were both still functional. Deletion of *rrp1* and *rrp2* was not found to have an effect on cell growth or sensitivity to DNA damage in previous studies (Dziadkowiec *et al.,* 2009) but mutants in which both *rrp1* or *rrp2* and *rad57* have been deleted were found to show greater sensitivity to DNA damage than Δ*rad57* alone. This was thought to be due to the fact that two different Rad51-dependent HR pathways exist, one which involves Rad55/57 and one which involves Swi5/Sfr1 and Rrp1 and Rrp2 (Dziadkowiec *et al.*, 2009). To test the function of the mEos3.1 tagged strains, survival assays were performed comparing *rrp1/2*-mEos3.1 and *rrp1*/2-mEos3.1 Δ*rad57* with Δ*rrp1*/2, Δ*rrp1*/2 Δ*rad57*, Δ*rad57* alone in the background used for the base strain creation and Δ*rad51*, in which the whole Rad51 dependent HR pathway is knocked out.

If the mEos3.1 tag was compromising the function of the gene it was expected that an increase in DNA damage sensitivity would be seen in the tagged strain in a Δ*rad57* background. The Δ*rad57* strains were created by amplifying the nourseothricin (Nat) resistance gene in the Mx6 plasmid with long arms of homology up and downstream of *rad57* and transforming into the various genetic backgrounds required and then selecting for resistance to Nat. On plates containing the DNA replication inhibitor HU the double delete strains (*rrp1*::*ura4 rad57*::*nat* and *rrp2*::*ura4 rad57*::*nat*) show greater sensitivity than the *rad57* single delete and similar sensitivity to the *rad51* delete strain (Figure 5.6). In contrast the tagged strains show the same sensitivity as the Δ*rad57* single delete and are less sensitive than the Δ*rad51*. A similar effect is seen when MMS, an alkylating agent, which causes damage in S phase, is used instead. These results indicate that the Rrp1 and Rrp2 function in HR is not significantly impaired by the mEos3.1 tag.

Figure 5.6 Assay of viability of mEos3.1 tagged strains

Each strain was grown overnight in liquid YE media to an OD_{595} of ~1. Serial dilutions were then performed and 7 µl of each dilution were pipetted onto YEA agar plates. The YEA plate with no additives was a positive control for loading. The other plates contained increasing amounts of the replication inhibitor HU and the alkylating agent MMS. AW310 is a wild type strain. On the plates containing DNA damaging agents it can be seen that AW310 and the *rrp1* and *rrp2* tagged strains were equally sensitive to DNA damage. The Δ*rad57* strain showed a little more sensitivity and the double mutants containing either Δ*rrp1* or Δ*rrp2* as well as Δ*rad57* deleted showed sensitivity almost equivalent to the Δ*rad5*1 mutant. The level of sensitivity shown by the strains containing tagged *rrp1* or *rrp2* with Δ*rad57* were most similar to that of Δ*rad57* alone suggesting the mEos3.1 tag is not strongly impacting on the function of Rrp1 or Rrp2.


YEA
AW310
rrp2 mEos .1
rrp1 mEos3.1
Δrhp57
rrp2::ura4 Δrhp57
rrp1::ura4 Δrhp57 1
rrp1::ura4 Δrhp57 2
rrp2 mEos3.1 Δrhp57 1
rrp2 mEos3.1 Δrhp57 2
rrp1 mEos3.1 Δrhp57 1
rrp1 mEos3.1 Δrhp57 2
Δrhp51
1mM HU
2 mM HU
3mM HU
AW310
rrp2 mEos .1
rrp1 mEos3.1
Δrhp57
rrp2::ura4 Δrhp57
rrp1::ura4 Δrhp57 1
rrp1::ura4 Δrhp57 2
rrp2 mEos3.1 Δrhp57 1
rrp2 mEos3.1 Δrhp57 2
rrp1 mEos3.1 Δrhp57 1
rrp1 mEos3.1 Δrhp57 2
Δrhp51
0.001% MMS
0.0015% MMS
AW310
rrp2 mEos .1
rrp1 mEos3.1
Δrhp57
rrp2::ura4 Δrhp57
rrp2 mEos3.1 Δrhp57 1
rrp2 mEos3.1 Δrhp57 2
rrp1 mEos3.1 Δrhp57 1
rrp1 mEos3.1 Δrhp57 2
Δrhp51

5.2.3 Rrp2 foci after DNA damage in fixed cells

Having established that the mEos3.1 tag did not appear to have an effect on the function of Rrp2 it was imaged using the PALM system. It has been reported that the proteins Rrp1 and Rrp2 form foci (clusters of many molecules) in the nucleus after induction of DNA damage with MMS treatment (Dziadkowiec *et al.,* 2009). However, these experiments were done using Rrp1/Rrp2 over-expressed from a plasmid rather than at endogenous levels. This was because the endogenous levels of these proteins are so low they were not detectable by a conventional fluorescence microscope. It was reasoned that, using a super-resolution microscope, it would be possible to visualise Rrp2 tagged C-terminally with mEos3.1 at endogenous levels. To replicate the results from Dziadkowiec *et al*., cells were treated with 0.1% MMS for 1 hr prior to imaging and compared to untreated cells. All cells were fixed using 1% formaldehyde in PBS.

In PALM microscopy super-resolution is achieved by separating fluorescent molecules by time and using a computer algorithm to calculate a localisation position for each molecule, which can then be plotted to give a super-resolution image. Localisations were seen in the nuclei of both the treated and untreated cells (Figure 5.7), indicating the presence of Rrp2 in the nucleus. However, foci (clusters of multiple molecules) were not observed and the distribution of localisations appeared to be the same in both treated and untreated cells. Next, the data were analysed using the ImageJ plugins PC-PALM Molecules and PC-PALM Clustering (see 3.1.1 for a thorough description). The PC-PALM Molecules plugin traces the localisations to remove repeat localisations and blinking effects then plots a binary image of localisations. The PC-PALM Clustering plugin then assigns localisations to clusters and plots a histogram of the size of these clusters. The algorithm used was particle centroid linkage with a 50 nm distance parameter, as described in Chapter 3. This was done for a sample of different nuclei from MMS treated and untreated cells and the resulting

frequency data were averaged and plotted as a histogram with error bars of 1 standard deviation. The algorithm has assigned clusters containing as many as 7 or 8 localisations, suggesting that there is some clustering of *rrp2*. However, there is no clear difference between the treated and untreated cells. More specifically there is no increase in clusters with a size >2 in the MMS treated cells.

Figure 5.7 Assessing Rrp2 foci in fixed cells after MMS damage

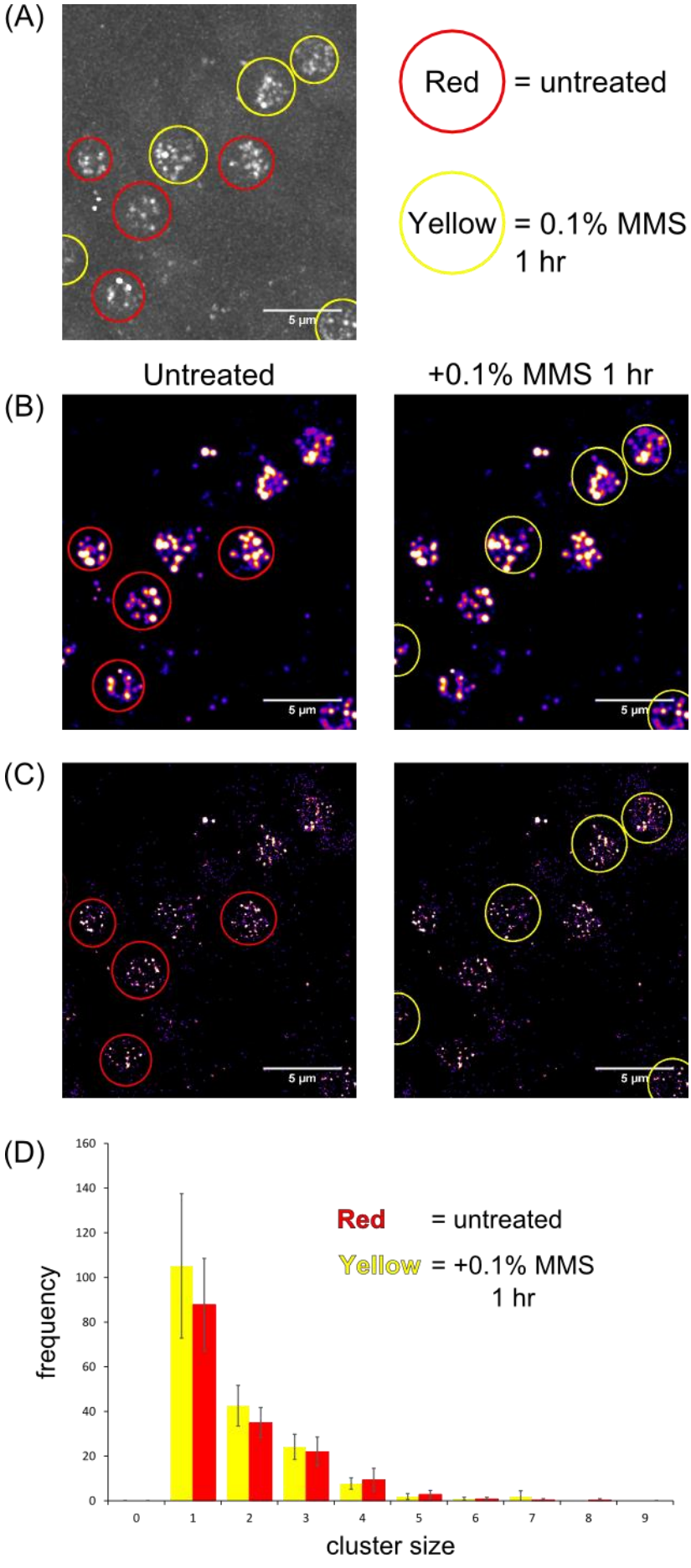
Cells were either treated with 0.1% MMS for 1 hr or grown untreated for 1 hr. Untreated cells were stained with the ConA dye conjugate before all of the cells were fixed for 30 minutes using 1% formaldehyde in PBS. Red circles indicate nuclei in untreated cells, yellow circles indicate nuclei in the MMS treated cells.

(A) The maximum projection of the dataset shows where the localisations are concentrated in the nucleus.

(B) The diffraction limited image shows that there are areas that could be clusters (the white and yellow areas have multiple localisations), but this could be a result of overlapping PSFs rather than physical clustering. No difference was seen between treated and untreated cells

(C) The super-resolution image also shows areas of multiple localisations. The improved resolution of this image suggests the lack of difference between treated and untreated cells is not an artefact of the diffraction limit.

(D) A histogram of the clustering of *rrp2* in the nuclei of fixed cells as calculated using the PCPALM Molecules and PCPALM Clusters plugins. Data are an average of 7 different nuclei for each treatment. Red = untreated, yellow = +0.1% MMS for 1 hour. There is no difference between the MMS treated and untreated cells in numbers of large clusters. Error bars = +/- 1 standard deviation.


(A)
Red = untreated
Yellow = 0.1% MMS 1 hr
5 µm
Untreated
+0.1% MMS 1 hr
(B)
5 µm
5 µm
(C)
5 µm
5 µm
(D)
Red = untreated
Yellow = +0.1% MMS 1 hr
frequency
cluster size
0
20
40
60
80
100
120
140
160
0
1
2
3
4
5
6
7
8
9

5.2.4 Cells with mEos3.1 tagged Rad52 show an increase in static localisations

The Rad52 protein is required during HR for targeting Rad51 to the ssDNA ends, allowing it to form a nucleoprotein filament and search for homologous sequences. When fluorescently tagged and expressed at endogenous levels, Rad52 has been shown to form foci (clusters of multiple molecules) on DNA, after DNA damage. It was expected that, using the motion blur technique, an increase would be seen in the number of localisations, indicating DNA-associated molecules, in response to DNA damage. As it is expected that Rrp2 is also forming foci on DNA in response to damage Rad52 is a useful positive control. Cells containing Rad52 tagged with mEos3.1 had previously been studied in the lab by Dr Thomas Etheridge. He had synchronised cells using a lactose gradient and treated these cells for 2.5 hrs, either with or without 0.005% MMS in YEA. The treated and untreated cells were imaged separately using the motion blur protocol. As expected the MMS treated cells showed a significantly higher number of localisations per nucleus than the untreated cells (Figure 5.8).

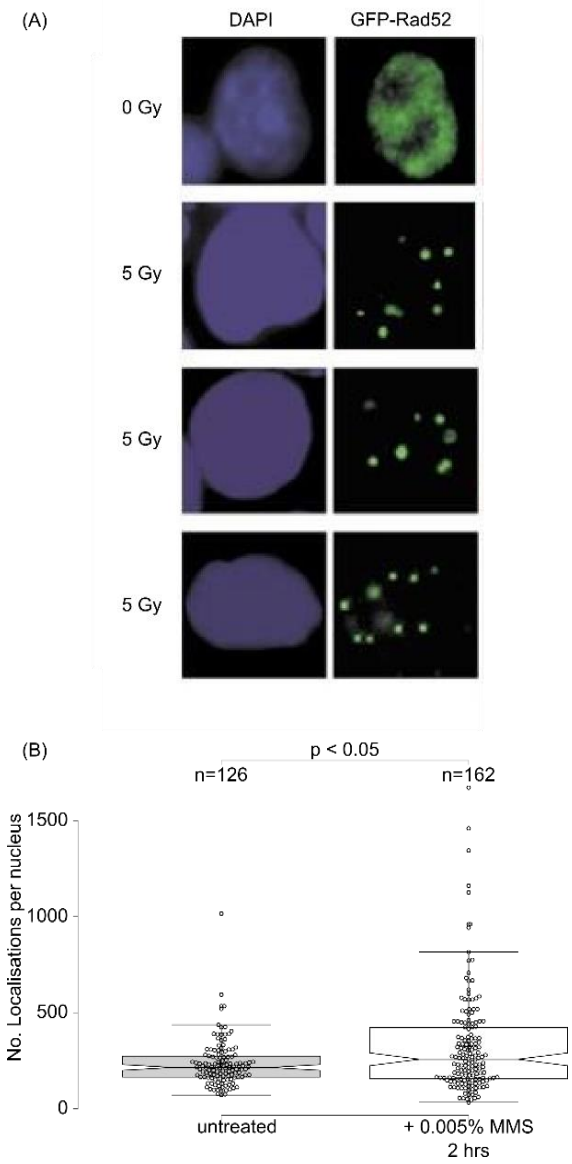


Figure 5.8 Rad52 tagged with mEos shows an increase in DNA bound portion in response to MMS

(A) Taken from Liu *et al.,* (1999). Widefield microscopy image of Rad52 foci forming after DNA damage induced by gamma radiation. The Rad52 image goes from diffusely spread fluorescence throughout the nucleus to localised fluorescence in a small number of tight foci after induction of DNA damage.

(B) Analysis of motion blurring PALM data (data were collected by Dr Thomas Etheridge) using mEos3.2 tagged Rad52. The number of localisations per nucleus was analysed and MMS treated and untreated plotted as box plots. It can be seen that the amount of DNA bound Rad52 appears to increase as a result of MMS treatment.

5.2.5 DNA-associated levels of Rrp2 are reduced after MMS damage

Having determined that it is not possible to see obvious Rrp2 focus formation in fixed cells it was decided to look at the DNA-association in live cells using motion blur PALM. The motion blur imaging technique can be used to visualise only those proteins that are close to stationary at the time of imaging. In the case of nuclear localised proteins, stationary can be assumed to mean associated with the DNA or chromatin. It was predicted that, if Rrp2 was localising to foci on the DNA in response to DNA damage, an increase in the number of localisations recalled by motion blur would be observed as a result of treatment with a DNA damaging agent. Asynchronous, log phase cells were either MMS-treated or left untreated. Samples were prepared and imaged separately. Unexpectedly a decrease was seen between the number of localisations seen in untreated cells and the number of localisations seen in MMS treated cells (Figure 5.9). To rule out the effects of preparing and imaging the MMS treated and untreated samples separately, the two different samples were imaged on the same slide. The origin of each cell could be traced by staining one sample with a Concavalin A (ConA) Alexa Fluor® 647 conjugate (see Materials and Methods: 2.3.5). These experiments also showed a decrease in localisation number in the MMS treated population. A single simultaneous imaging experiment consisting of 6 separate imaging sets was taken over the course of roughly 2.5 hours. To establish whether the number of localisations seen was changing significantly over the experimental time frame the data from each imaging set was plotted separately in the order of acquisition. No particular trend was seen from plotting the data in this way (Figure 5.9).

Figure 5.9 DNA-associated levels of Rrp2 are reduced after MMS treatment

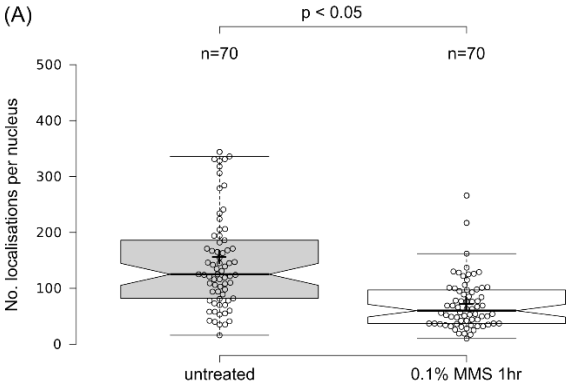
(A) Asynchronous populations of MMS treated and untreated cells were imaged separately using the motion blur PALM technique. The number of localisations per nucleus was counted for all in-focus nuclei and plotted as a box plot. The number of localisations was reduced in the MMS damaged cells which is in contrast to the widefield imaging of overexpressed Rrp2 (Dziadkowiec *et al.*, 2013).

(B) To rule out effects from preparing the MMS treated and untreated cells separately cells were imaged simultaneously by staining one set of cells with a ConA dye conjugate which binds the surface of cells. As above a decrease in number of localisations was seen after MMS treatment.

(C) A time course of simultaneous imaging data collected in six separate image sets over the course of approximately 2.5 hours. Data were plotted separately by image set in the order of acquisition. No obvious change in the number of localisations was seen over time. At all time points there were fewer localisations in the MMS treated cells.

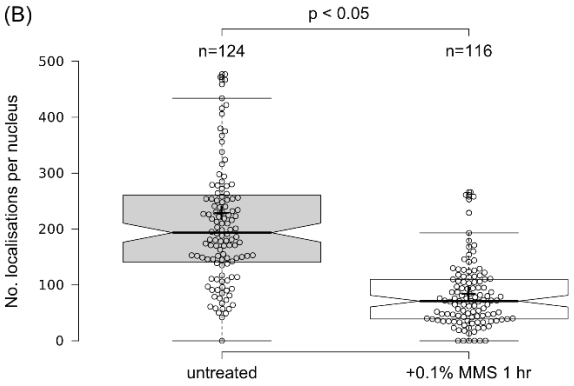
Rrp2 mEos3.1 localisations

(A)



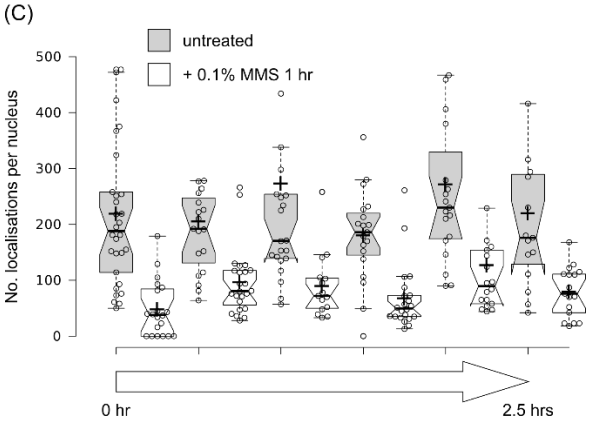
Rrp2 mEos3.1 localisations with simultaneous imaging

(B)



Rrp2 localisations split over time

(C)



5.2.6 Bleed through from the Concavalin A Alexa Fluor dye is not affecting localisation number

When imaging the ConA AlexaFluor® 647 conjugate can bleed through into the 561 imaging channel. To ensure that this bleed through was not affecting the number of localisations seen, a sample of *rrp2-mEos3.1* cells were stained with ConA dye and a sample were left unstained, then both were imaged simultaneously. After analysis, both showed the same distribution of number of localisations suggesting that any bleed through that was happening was not strongly affecting the analysis of localisation numbers (Figure 5.10).

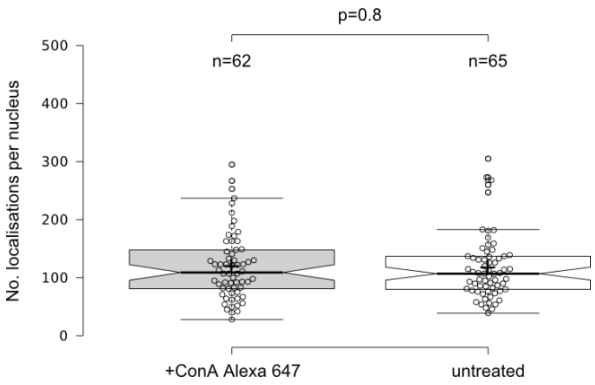


Figure 5.10 ConA AlexaFluor® 647 bleed-through does not affect number of localisations seen

To ensure that bleed-through into the 561 channel from the ConA dye conjugate was not affecting the localisation numbers wild type Rrp2 mEos3.1 cells were imaged after half were treated with ConA and half not. Both populations show a similar distribution of number of localisations with no significant difference.

5.2.7 Cell cycle profile of Rrp2 localisation

To determine whether the differences in localisation numbers were due to cell cycle profile cells were synchronised using lactose gradients to select G2 cells. These cells were then allowed to grow in the presence or absence of MMS for 2.5 hrs. After 2.5 hrs the majority of cells were in S phase as judged by septation index. In *S. pombe* G1 is very short and S phase is co-incident with cytokinesis, which can be judged by septation. This means S phase cells can be distinguished as bi-nucleate cells from mono-nucleate G2 cells facilitating a direct comparison of S and G2 cells.

Since 0.1% MMS is considered quite a high concentration when compared to many protocols, and there was the possibility that it might be quenching or otherwise inhibiting the fluorescence in some way, it was decided to to perform some experiments using the lower concentration of 0.005% MMS. The higher dose was used on other occasions to maintain consistency with the work published by the Dziadkowiec group. Figure 5.11 A shows that even with this low dose of MMS the number of Rrp2 localisations was still reduced.

Figure 5.11 A reduction is seen in DNA-associated Rrp2 in cultures enriched for S phase cells

G2 cells were separated by size on a lactose gradient and grown on for 2.5 hours to S phase. Half of the cells were grown in the presence of 0.005% MMS and the other half were untreated. The untreated cells were stained with the ConA AlexFluor® conjugate and both treated and untreated were imaged simultaneously using the motion blurring protocol

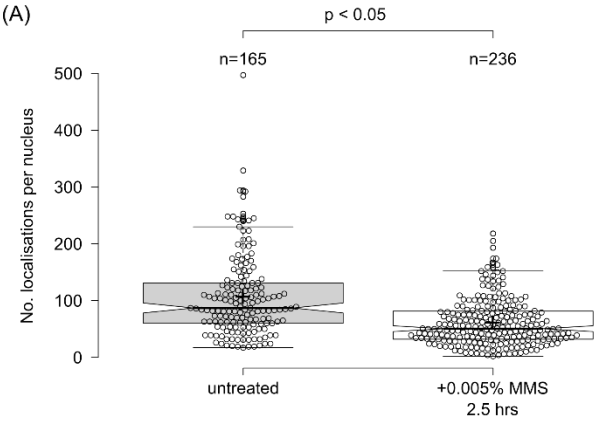
(A) Total data shows that a lower dose of 0.005% MMS still leads to a reduction in localisation numbers compared to untreated cells.

(B) Untreated bi-nucleate (S phase) cells have fewer localisations than untreated mono-nucleate (G2) cells.

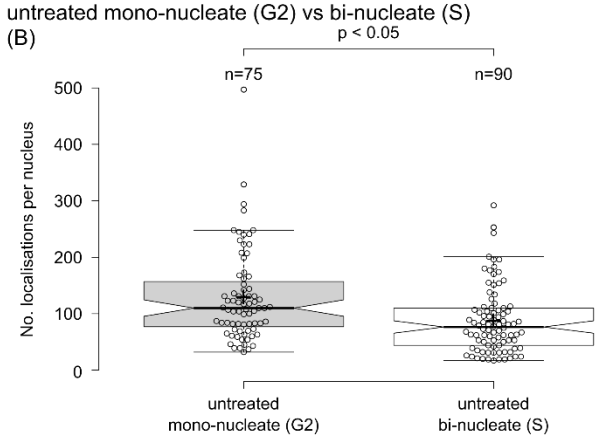
(C) The bi-nucleate cells show a reduction in number of localisations in response to treatment with MMS.

(D) The difference in localisation numbers is less pronounced between mono-nucleate and bi-nucleate cells after MMS treatment. Box plots for bi-nucleate (S phase) cells are coloured blue.

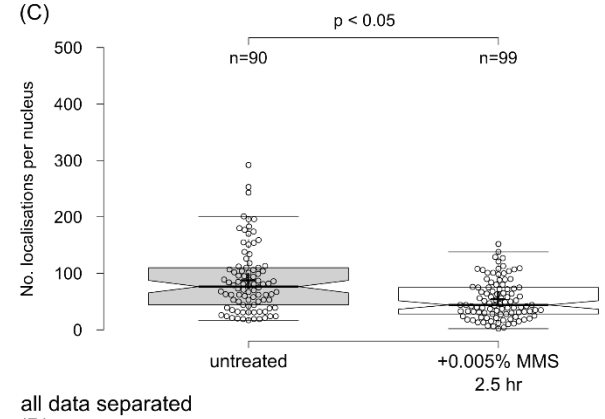
synchronised cells - bulk data



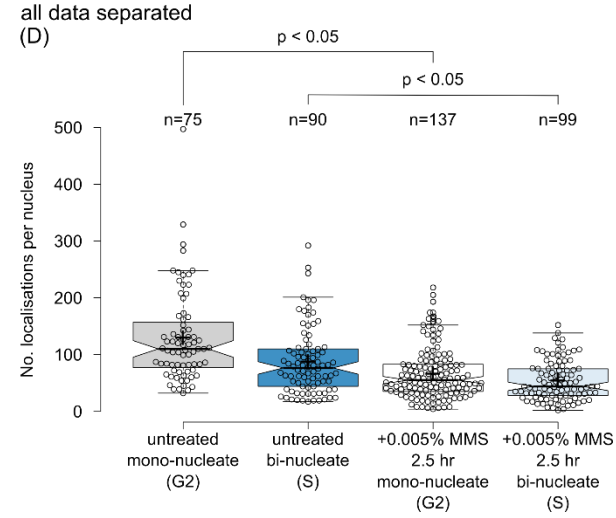
untreated mono-nucleate (G2) vs bi-nucleate (S)



bi-nucleate (S) only



all data separated



There was a difference seen in levels of stationary Rrp2 between bi-nucleate (S phase) and mono-nucleate (G2) cells, with bi-nucleate cells having fewer localisations (Figure 5.11 B). One possible reason for this is that the Rrp2 content has been split between the two nuclei by mitosis. Data published by the Bähler group suggest that the expression profile of Rrp2 does not vary across the cell cycle (Rustici *et al.*, 2004) but without knowing the stability of the protein this doesn't give much information about the overall protein content as a function of the cell cycle. However there was still a decrease in MMS treated bi-nucleate cells compared to untreated bi-nucleate cells (Figure 5.11 C). The difference between bi-nucleate and mono-nucleate was less pronounced in the MMS treated cells (Figure 5.11 D). This suggests that the reduction in the bound fraction seen as a result of MMS treatment is not specific to a particular phase of the cell cycle.

5.2.8 The change in localisation numbers is not an MMS-specific effect

To determine whether the reduction in localisations was linked to DNA damage in general, as opposed to being an MMS-specific effect, other agents were examined. Cells blocked in S phase by treatment with HU can efficiently re-start replication on removal of HU in a manner that is thought to require the stablilisation of the arrested fork by a process that involves HR proteins. HR is also required to restart a subset of collapsed, or damaged, replication forks. This later process has been shown to be temporally separated from replication (Meister *et al.*, 2005). Cells were synchronised in G2 and grown for 2.5 hrs in the presence of 10 mM HU. As before, untreated cells were labelled with the ConA conjugate. A significant decrease in localisation number was seen in cells treated with HU (arrested in S phase) compared to untreated, bi-nucleate (S phase) cells (Figure 5.12 A).

In cells with fluorescently tagged Rad52 an increase in fluorescent foci is not observed until after release from HU treatment (Meister *et al*., 2005; Sabatinos *et al*., 2012), consistent with HU induced DNA damage not being repaired immediately. To see whether a related inverse effect would be observed for Rrp2, cells were synchronised in G2 and grown for 2.5 hrs in the presence of 10 mM HU. The HU was then washed out and the cells returned to growth media with time points taken at 0, 30 and 45 minutes. Cells from 30 and 45 minutes after HU release were imaged simultaneously with the 0 time point cells. The results show a decrease in localisations from time point 0 to 30 and 45 minutes (Figure 5.12 B).

HU is thought to only affect S phase cells and MMS has its genotoxic consequences when cells pass through S phase. To compare a different type of damaging agent that is known to cause damage in G2 cells, asynchronous mEos3.1 tagged cells were irradiated with 250 Gy of IR then allowed to recover for 30 minutes. As ~70% of asynchronous *S. pombe* cells are in G2 and, after irradiation with IR, would delay mitosis and remain in G2, analysis of mono-nucleate cells would largely reflect G2 cells that were irradiated and remained in G2. Again, the treated cells show fewer localisations per nucleus than the untreated cells (Figure 5.12 C). These observations suggest that the reduction in the DNA-associated population is not an MMS-specific effect and support the hypothesis that it is not specific to a cell cycle phase.

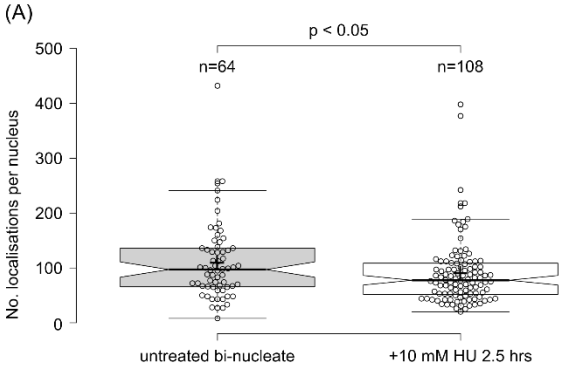
Figure 5.12 A decrease in DNA-associated Rrp2 is seen in response to HU and Gamma radiation

(A) Cells were synchronised in G2 and grown with 10 mM HU for 2.5 hours. The number of localisations was reduced in the HU-arrested cells compared to unperturbed S phase (bi-nucleate) cells.

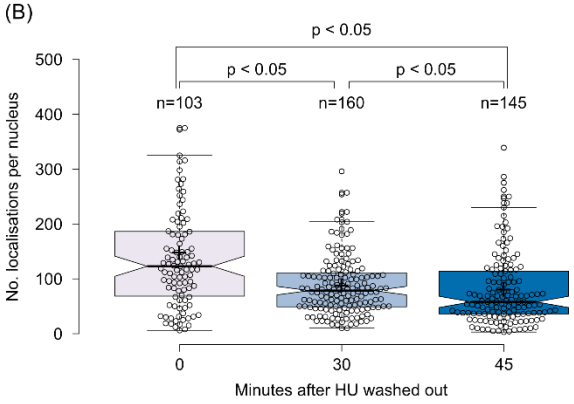
(B) Cells were HU treated for 2.5 hrs then allowed to recover in YE. Samples were collected after 0, 30 and 45 minutes of recovery time. Two sets of imaging were carried out, one with 0 and 30 minute time points imaged simultaneously and one with 0 and 45 minute time points. Results show a decrease from the 0 time point to the 30 minutes time point and a further slight decrease to the 45 minutes time point.

(C) Asynchronous cells were exposed to 250 Gy of gamma radiation and then allowed to recover in YE for 30 minutes. A significant reduction in localisation numbers was observed.

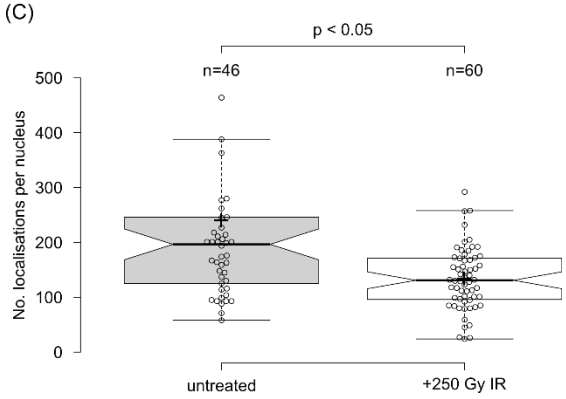
Rrp2 localisations with HU
(A)



Rrp2 localisations after washing out HU
(B)



Rrp2 localisations after ionising radiation
(C)



5.2.9 Rad57 deficient mutants show displacement of Rrp2 after DNA damage

To explain this reduction in the DNA-associated portion of Rrp2 after damage one hypothesis considered was that, upon DNA damage, the Rad55/57 pathway is used preferentially to the Swi5/Sfr1 pathway and therefore Rrp2 is displaced from the DNA by Rad55/57. This was tested by imaging a strain in which *rad57* had been replaced with the *nat* resistance gene. Asynchronous cells were treated with 0.1% MMS for one hour and untreated and MMS treated cells were imaged simultaneously, with untreated cells stained with the ConA conjugate. A significant decrease in the number of localisations in the MMS treated cells compared to the untreated cells was observed (Figure 5.13 A) but it appeared to be less than that in the *rad57*$^{+}$ cells. It can be concluded from this that Rad57 is not required in order for Rp2 to disassociate from the DNA in the event of DNA damage.

To see whether the deletion of *rad57* affected the amount of Rrp2 bound to the DNA both with and without damage untreated *rad57*$^{+}$and Δ*rad57* cells were imaged simultaneously and then MMS treated *rad57*$^{+}$ and Δ*rad57* cells were imaged simultaneously (Figure 5.13 B). From this it can be seen that the Δ*rad57* cells have fewer Rrp2 localisations before MMS treatment than *rad57*$^{+}$ cells do. After DNA damage the levels show no difference between the Δ*rad57* mutant and the *rad57*$^{+}$. Cells containing the Δ*rad57* mutation are prone to the accumulation of spontaneous DNA damage (they are less efficient at repairing spontaneous lesion, although we cannot rule out they are also generating increased numbers of spontaneous lesions) so a potential explanation for this observation is that the Rad57 deficient cells show less DNA bound Rrp2 because they had already accumulated spontaneous DNA damage.

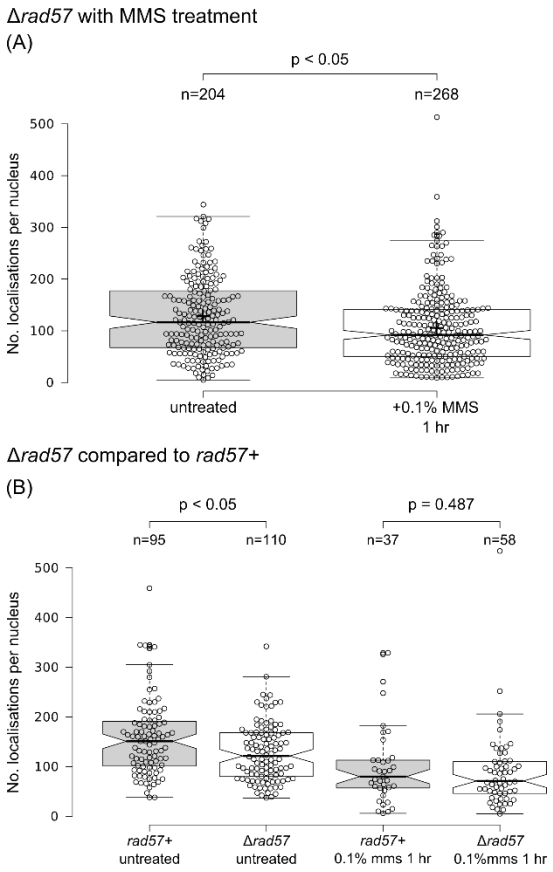


Figure 5.13 A decrease in DNA-associated Rrp2 is seen in *rad57* delete backgrounds

(A) Δ*rad57* cells were imaged after MMS treatment. A decrease in the number of localisations per nucleus was seen in cells that had been treated with 0.1% MMS for 1 hr compared to untreated. This suggests that the absence of Rad57 does not affect the ability of Rrp2 to unbind from DNA.

(B) To examine whether there was any difference between the Δ*rad57* cells and the *rad57*$^{+}$ cells untreated *rad57*$^{+}$ and Δ*rad57* cells were imaged simultaneously and then MMS treated *rad57*$^{+}$ and Δ*rad57* cells were imaged simultaneously. The results suggest that the untreated Δ*rad57* cells have less DNA-associated Rrp2 than the *rad57*$^{+}$ in the absence of DNA damage, but both cell types have similar levels of DNA bound Rrp2 after treatment with MMS.

5.2.10 The decrease in DNA-associated Rrp2 is not brought about by the DNA damage or intra-S phase checkpoint responses

Many of the processes that occur in response to DNA damage are triggered by one of the cell cycle checkpoint pathways. It is thus possible that the disassociation of Rrp2 from the DNA in response to DNA damage is a consequence of checkpoint pathways being activated (See Introduction: 1.3.3.3). The checkpoint pathways ensure that cells in which DNA has been damaged or replication has halted do not progress further through the cell cycle and also coordinate a wide range of DNA damage responses, such as activation of DNA repair pathways. Checkpoint pathways can be activated at three different points in the cell cycle; G1/S, preventing the cell from beginning replication of damaged chromosomes, intra-S which slows down DNA replication and regulates DNA repair pathways and G2/M which inhibits transition from G2 into mitosis and regulates DNA repair (Furuya and Carr, 2003).

A number of genes have been documented as being involved in the DNA damage checkpoint. In *S. pombe* the protein Rad3 (mammalian ATR) is a phosphoinositide 3-kinase related kinase involved in sensing DNA damage and activating the checkpoint (Bentley *et al.*, 1996). When activated, one of the roles of Rad3 is to activate the effector kinases Chk1 or Cds1 (Walworth and Bernards, 1996; Lindsay *et al.*, 1998) and to delay cell cycle progression through inhibition of Cdc2 (Boddy *et al.*, 1998; Furnari *et al.*, 1999), the cyclin dependent kinase responsible for progression into mitosis (Gould and Nurse, 1989). In *S. pombe* none of the core checkpoint genes is essential, but deletion leads to increased DNA damage sensitivity. Since the purpose of the checkpoint is to prevent cell cycle progression when there is DNA damage its activity also regulates the HR repair process (reviewed in Kara A. Nyberg *et al.*, 2002; Zhou and Elledge, 2000). If the removal of Rrp2 from the

DNA were triggered by activation of the checkpoint pathway then it would be expected that, in checkpoint deficient cells, a decrease in localisation numbers as a result of DNA damage would not be seen.

Cells expressing mEos3.1 tagged *rrp2* were crossed with various checkpoint defective mutants; *rad3*::*kan* and *cds1*::*nat* were provided by Dr Thomas Etheridge (University of Sussex), *chk1*::*kan* came from the strain collection of Dr Ellen Tsang (formerly University of Sussex). For each of these genetic backgrounds untreated cells were compared with cells treated with 0.1% MMS for 1 hour. In each case the number of localisations seen decreased after MMS treatment (Figure 5.14 A, C & E).

Some of the checkpoint null mutants appeared to show lower levels of DNA bound Rrp2 than the checkpoint positive strain in the untreated condition. To check this experiments were performed where each deletion mutant was imaged, untreated, simultaneously with untreated cells lacking the deletion (Figure 5.14 B, D & F). In these experiments no difference was seen between the null mutant backgrounds and the positive strains. As such any difference observed in localisation numbers in untreated cells can therefore be attributed to natural variation between experiments, possibly resulting from variation in background fluorescence or the stochastic nature of the photo-activation process and its efficiency. These observations suggest that the reduction in DNA-associated Rrp2 as a result of DNA damage is not directly triggered by checkpoint activation.

Figure 5.14 A decrease in Rrp2 associated with DNA is seen in various checkpoint gene deletion backgrounds

(A) In the Δ*chk1* background there was a significant decrease in the number of localisations seen after MMS treatment compared to untreated.

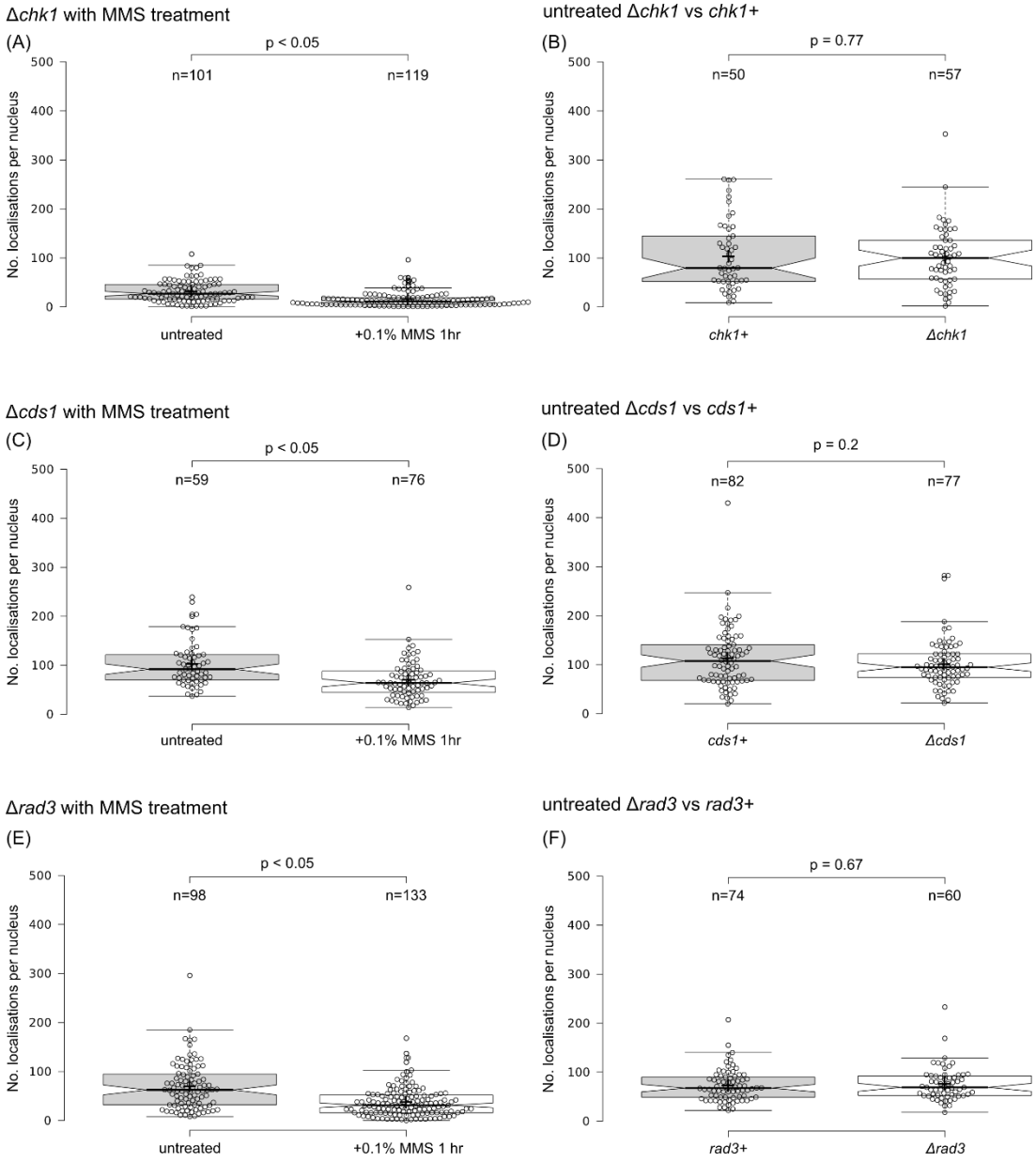
(B) Comparing *chk1*$^{+}$ untreated cells to Δ*chk1* untreated cells no significant difference was seen.

(C) In the Δ*cds1* background there was a significant decrease in the number of localisations seen after MMS treatment compared to untreated.

(D) Comparing *cds1*$^{+}$ untreated cells to Δ*cds1* untreated cells no significant difference was seen.

(E) In the Δ*rad3* background there was a significant decrease in the number of localisations seen after MMS treatment compared to untreated.

(F) Comparing *rad3*$^{+}$ untreated cells to Δ*rad3* untreated cells no significant difference was seen.


Δchk1 with MMS treatment
(A)
p < 0.05
n=101
n=119
No. localisations per nucleus
untreated
+0.1% MMS 1hr
untreated Δchk1 vs chk1+
(B)
p = 0.77
n=50
n=57
chk1+
Δchk1
Δcds1 with MMS treatment
(C)
p < 0.05
n=59
n=76
untreated
+0.1% MMS 1hr
untreated Δcds1 vs cds1+
(D)
p = 0.2
n=82
n=77
cds1+
Δcds1
Δrad3 with MMS treatment
(E)
p < 0.05
n=98
n=133
untreated
+0.1% MMS 1 hr
untreated Δrad3 vs rad3+
(F)
p = 0.67
n=74
n=60
rad3+
Δrad3

5.2.11 The decrease in DNA-associated Rrp2 is not dependent on heterochromatin

Heterochromatin consists of highly compacted, highly ordered regions of the DNA. In contrast, euchromatin is less dense and generally more easily transcribed (reviewed in Grewal and Jia, 2007). Heterochromatin is mainly found at areas of highly repetitive DNA (e.g. around centromeres and telomeres) (reviewed in Elgin and Grewal, 2003) or at developmentally regulated genes (Lu *et al.*, 1998). Compaction into heterochromatin is used to control transcription (reviewed in Cremer and Cremer, 2001), recombination (Bisht *et al.*, 2008) and DNA repair (reviewed in Cann and Dellaire, 2010). In *S. pombe* Clr4 (homolog of human SUV39H1), a histone methyltransferase which methylates histone H3 at the lys9 residue (Rea *et al.*, 2000), and the Swi6 protein (homolog of mammalian HP1) are both required for the stable formation of heterochromatin (Cheutin *et al.*, 2003; Peng and Karpen, 2009). Swi6 binds to the methylated histone H3-K9 (Bannister *et al.*, 2001) and forms an assembly platform for the other proteins involved in heterochromatin stabilisation and spreading (Fischer *et al.*, 2009; Zofall and Grewal, 2006).

Dziadkowiec *et al.*, found that in a multiple sequence alignment with Rad8 and Ris1, Rrp1 and Rrp2 scored highly on their SNF2-N, Helicase-C and RING finger domains (Dziadkowiec *et al.*, 2013) suggesting that they could belong to group 7 of the SNF2 family, a family of chromatin remodelling proteins with a common SNF2 domain (reviewed in Ryan and Owen-Hughes, 2011). This being the case, Rrp2 might need the methylation of histone H3 and Swi6 in order to assemble onto the DNA in undamaged cells and, in the absence of either, it would then be seen that Rrp2 is not associated with the DNA in the undamaged cells and association would not decrease in response to DNA damage.

To assess whether the change in DNA-associated Rrp2 with and without damage was related to heterochromatin an *rrp2-mEos3.1* containing strain was crossed into two strains

deficient in heterochromatin. The *swi6*::*ura4* strain was provided by Dr Edgar Hartsuiker (Bangor University) and the *clr4*::*nat* strain was provided by Dr Jo Murray (University of Sussex). As before, cells were treated with 0.1% MMS for 1 hour and imaged simultaneously with untreated cells that were labelled with the ConA dye conjugate. In both backgrounds a similar number of localisations was seen in undamaged cells to previous experiments and there was a decrease in number of localisations from untreated to MMS treated cells (Figure 5.15). This suggests that heterochromatin structure does not play a role in how Rrp2 binds to the DNA.

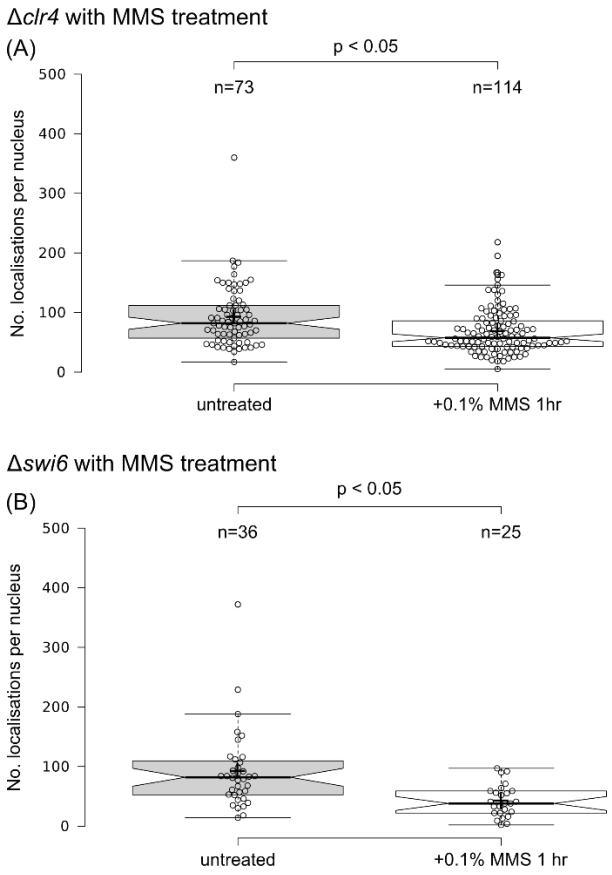


Figure 5.15 Heterochromatin deficiency has no effect on Rrp2 association with DNA

Heterochromatin deficient cells were either treated with 0.1% MMS for 1 hr or untreated.

(A) In cells deficient in the histone methyltransferase Clr4 (Δ*clr4*) the number of localisations seen per nucleus was reduced in the MMS treated cells, similarly to wild type cells

(B) In cells deficient in the heterochromatin protein Swi6 (Δ*swi6*) the number of localisations, and hence the level of Rrp2 associated with the DNA, was reduced by MMS treatment, similarly to wild type

5.2.12 Imaging Rrp1 N-terminally tagged with mEos3.1

The data from yeast two-hybrid studies from the Dziadkowiec lab suggest that Rrp1 and Rrp2 interact (Dziadkowiec *et al.*, 2013) leading to the hypothesis that they function together as a heterodimer. The Rrp1 N-terminal mEos3.1 strain was created (as detailed in 5.2.1.2). Sequencing of the strains suggested that the mEos3.1 sequence was present in the correct position and in frame with *rrp1* (see Appendix for sequencing results). However, it was not possible to detect Rrp1 under the microscope. Activations were not clearly seen, in contrast to the Rrp2 experiments, and localisations were spread throughout the cell (Figure 5.16) rather than localised to the nucleus as expected. In all probability, and based on a by-eye assessment of the raw imaging data, these are localisations of autofluorescent noise in the cell. One possible conclusion is that there is even less Rrp1 in the cells than there is Rrp2. However the mEos3.1 tag can interfere with gene function (Carr lab, unpublished data). As such it is possible that mEos3.1 is being cleaved from the Rrp1 protein in some way, as was seen with $Cdc22^{R1}$ mEos3.2.

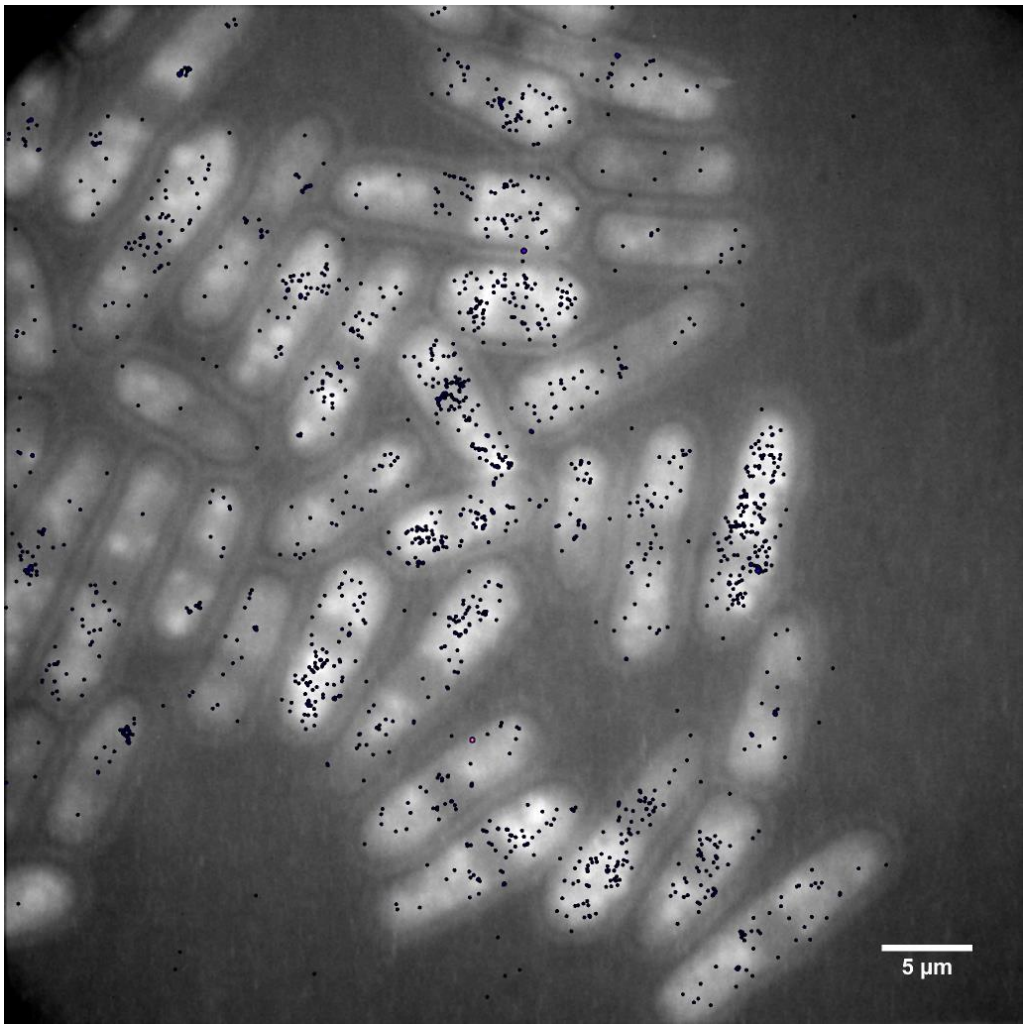


Figure 5.16 Imaging Rrp1 tagged N-terminally with mEos3.1

Live motion blur imaging of untreated Rrp1 N-term mEos3.1 cells. Localisation plots have been superimposed onto a sum of intensity projection to show cell location.

5.2.13 Rrp1 does not affect the disassociation of Rrp2 from DNA

It is assumed that Rrp1 and Rrp2 act together as a complex because of yeast two-hybrid assay interaction data, the fact that Rrp1 and Rrp2 foci appeared to co-localise when over-expressed and the observation that the triple *rrp1 rrp2 rad57* deleted strain is no more sensitive to DNA damage than the *rrp1 rad57* or the *rrp2 rad57* double mutants (Dziadkowiec et al., 2009, 2013). If Rrp1 and Rrp2 act as a heterodimer it would be expected that the loss of Rrp1 would affect the localisation of Rrp2. Looking at Rrp2-mEos3.1 in the *rrp1* delete background using the motion blur protocol the number of localisations seen after treatment with MMS were again lower than in untreated cells (Figure 5.17). However when untreated Δ*rrp1* cells were imaged with untreated *rrp1*$^{+}$ cells the localisation numbers were significantly higher in the *rrp1*$^{+}$ cells. This suggests that Rrp1 may be partially required for binding of Rrp2 to DNA in undamaged cells but that it is not needed for the removal of Rrp2 from damaged DNA.

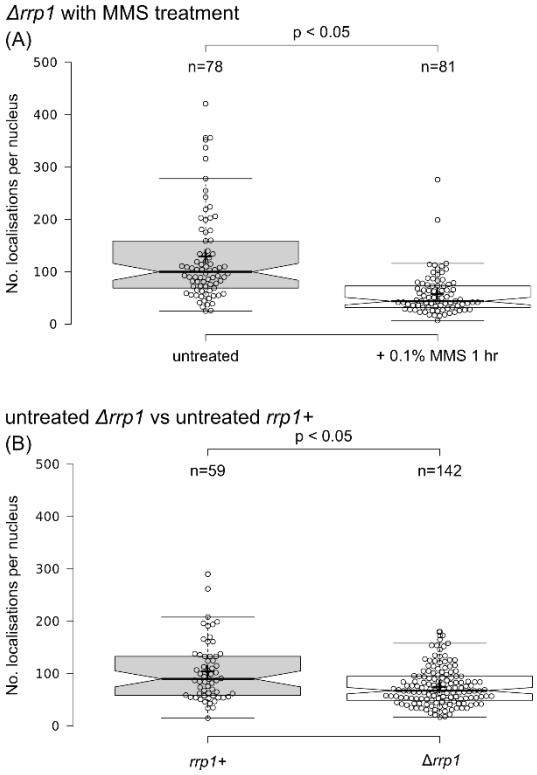


Figure 5.17 DNA bound Rrp2 decreases after damage in the *rrp1* delete strain

(A) In the Δ*rrp1* background the level of DNA bound Rrp2 decreases in MMS treated cells in comparison to untreated cells.

(B) Untreated Δ*rrp1* cells were imaged simultaneously with untreated *rrp1*$^{+}$ cells. The number of localisations in the untreated Δ*rrp1* cells is significantly lower than in the untreated *rrp1*$^{+}$cells.

5.3 Discussion

5.3.1 Rrp2 is associated with DNA in unperturbed cells

Rrp2 was C-terminally tagged with mEos3.1 using recombination mediated cassette exchange in order to use PALM to study it at endogenous levels. Previous studies had shown that when over-expressed, Rrp2 formed foci in the nucleus when cells were damaged with high doses of MMS. This foci formation was not observed using PALM and fixed cells treated with MMS. Using the PALM motion blur technique (Etheridge *et al.*, 2014), Rrp2 was imaged in live cells and a static fraction was observed in undamaged cells. However, the striking and unexpected observation is that this fraction was reduced after DNA damage at times consistent with HR. This is the opposite of what would be expected for an HR protein, for example Rad52. This is interesting because Dziadkowiec *et al.,* showed that Rrp2 was only forming foci on the DNA when there was damage (Dziadkowiec *et al.,* 2013) but since they were only able to observe over-expressed Rrp2 it may not reflect behaviour at endogenous levels. In this analysis, visualising Rrp2 at endogenous levels, no difference in distribution (i.e. evidence for foci induction) in fixed cells was seen between cells treated with MMS and cells left untreated. The fact that foci were not observed in fixed cells seems to be confirmed by the lack of increase in localisation numbers after damage treatment in the motion blur experiments.

Additional unpublished data from the Dziadkowiec lab has shown that when Rrp1 and Rrp2 are overexpressed from a vector they can sometimes form foci spontaneously. They also found that prolonged overexpression of either protein causes growth defects and observed nuclear morphology aberrations and chromosome loss. This suggests that visible foci could be an artefact of overexpression, which is why they were not seen in the PALM experiments. The facts that over-expression leads to these spontaneous foci formation, and appears to be damaging to the cell, and the observation that, upon DNA damage, an

amount of the Rrp2 becomes disassociated from the DNA suggests that perhaps it is deleterious for the cell to have too much Rrp2 when damage has occurred. One theory is that the Rad55/57 pathway of Rad51-dependent HR is somehow blocked by an excess of Rrp1 or Rrp2, limiting the cell's options after DNA damage.

Using ChIP the Dziadkowiec group have shown that, at endogenous levels, GFP-tagged Rrp1 is enriched at the site of induced DSBs (D. Dziadkowiec, personal communication). In addition they were able to Co-IP Rrp1 and Rad52 with MMS treatment but not without and Co-IP Rrp1 and Rrp2. These results strongly suggest that Rrp1 is going to sites of damage and that Rrp1 and Rrp2 associate. This does not contradict the data presented: potentially DNA-associated Rrp2 is still observed after damage, which may well be at sites of damage – albeit not in quantities that can be distinguished as foci.

5.3.2 Rrp2 disassociates from the DNA in response to damage

Fewer Rrp2 localisations were seen per nucleus after treatment with a DNA alkylating agent, which can cause DNA damage. This leads to the conclusion that Rrp2 is associated with the DNA under normal circumstances and that it is disassociated from the DNA in response to DNA damage. One possible reason for this is that it is performing another function on the DNA under normal conditions and that, upon damage, it comes off the DNA and only a proportion of it goes back onto the DNA at the damage sites. Interesting questions thrown up by this analysis are: what is Rrp2's function in undamaged cells? What is the mechanism by which it comes off DNA in response to DNA damage? What is its precise function in the Rad51-dependent pathway of HR?

5.3.3 Rrp2 removal coincides with the expected start of HR

Meister *et al.*, (2005) studied the timing of repair when cells are arrested in S phase due to nucleotide depletion. They showed that in wild type cells there is a transient accumulation of RPA followed by Rad52 at 0.5hr after release from HU, after replication has resumed. They suggested this represented a period of fork repair and recovery (Meister *et al.*, 2005; Sabatinos *et al.*, 2012). The number of Rrp2 localisations in HU arrested cells was shown to be reduced when compared to untreated cells and to reduce further after a recovery period of 30 minutes. The timing of this reduction coincides with when it is expected that HR would begin in cells released from HU arrest.

5.3.4 Checkpoint and heterochromatin mutants do not affect Rrp2's response to DNA damage

The ability to sense damage and activate a checkpoint to delay cell cycle progression and the ability to regulate damage repair machinery are linked. However, genetic deficiency in three major checkpoint pathway genes, *rad3* (ATR, required for both S and G2 checkpoints), *chk1* (required for G2) and *cds1* (required for the S phase checkpoint in *S. pombe*) did not seem to affect either the pre-damage level of DNA-associated Rrp2 or the post damage level of DNA-associated Rrp2. This suggests that, whatever mechanism is causing Rrp2 to come off the DNA after DNA damage, it is not checkpoint pathway regulated.

The SNF2-N domain of Rrp2 suggested that it might belong to a family of chromatin remodelling proteins and it was speculated that it might therefore require heterochromatin for association with DNA. In mutants deficient in two important heterochromatin genes, *clr4* (the histone methyltransferase which methylates histone H3) and *swi6* (which binds to

the methylated histone H3 and forms an assembly platform) there was no change in the behaviour of DNA-associated Rrp2 either before or after DNA damage treatment. If Rrp2 were being recruited by heterochromatin in undamaged cells it would be expected that fewer or no localisations would be seen in the undamaged, heterochromatin deficient cells. These results suggest Rrp2 binding to DNA is not regulated by heterochromatin. Further to this, if Rrp2 was mainly binding at heterochromatic regions it would be expected that a pattern of localisations at the periphery of the nucleus (where the majority of heterochromatin is located) would be observed in fixed cells, and this was not the case.

5.3.5 Rrp2 and Top2

Etoposide is an anti-cancer drug approved by the FDA in 1983. It acts by inhibiting type II topoisomerases (Top2s) (Wu *et al.*, 2011), which are required to manage topological problems in DNA such as supercoiling. *S. pombe* cells are resistant to etoposide-induced genotoxic stress but deletion of the drug efflux pump gene, *pmd1*, renders them sensitive (Arita *et al.*, 2011). The lab of Dr Li-Lin Du performed a genetic screen based on this discovery to look for increased etoposide sensitivity in a Δ*pmd1* background (L, Du, personal communication, 2016). The top hit for sensitivity in their screen was an *rrp2* deletion: when the *rrp2* gene was deleted in a Δ*pmd1* background they observed significantly increased growth inhibition upon treatment with etoposide. Using a cell survival assay to validate the screen results they showed that the double delete mutant (Δ*pmd1* Δ*rrp2*) was more sensitive to etoposide than the single mutant (Δ*pmd1*), but was not more sensitive to other DNA damaging agents such as HU and MMS. In a further screen for suppressor mutations that could rescue etoposide sensitivity in the double deletion mutant Du *et al.*, found that many of the hits they were getting were SUMOylation related enzymes. SUMO is a small ubiquitin-like modifier which has critical roles in DNA

transcription, DNA replication and the response to DNA damage (reviewed in Dou *et al.*, 2011). Rrp2 has multiple SUMO-interacting motifs in its N-terminal domain and Du *et al.*, showed that Rrp2 interacts with SUMOylated Top2 and that deletion of the SUMO interacting motifs reduces Rrp2's ability to confer resistance to etoposide. In Δ*rrp2* mutants they noted that the levels of SUMOylated and chromatin bound Top2 were reduced. As a result they hypothesised that Rrp2 binds to SUMOylated Top2 and prevents its degradation by SUMO targeted ubiquitin ligase (STUbL). This may mean that the changes seen in the static population of Rrp2 in the motion blur experiments are related to the SUMOylation state of Top2. As such, it would be very interesting to repeat the Rrp2 motion blur experiments in cell lines deficient for SUMOylation, for example by deletion of *pli1*, which is one of the genes found by Du *et al.*, in their screen for rescue of Δ*pmd1* Δ*rrp2* etoposide sensitivity. Alternatively the cassette exchange system could be used to create SUMO interacting motif deletion mutants of Rrp2 tagged with mEos3.1 for motion blur experiments.

Interestingly *rrp1* was not picked out by the original screen for etoposide sensitivity, and, when tested separately was not seen to cause etoposide sensitivity in the absence of *pmd1*. This suggests that the two proteins Rrp1 and Rrp2 may not be as functionally similar as initially assumed.

5.3.6 Further experiments

5.3.6.1 Tagging Rrp1 with a different fluorophore

The Rrp2 protein has been assumed to work in a complex with Rrp1, at least during HR repair. Dziadkowiec *et al.,* saw them co-localise (when over-expressed) and they both function in the same pathway of Rad51-dependent HR. It was not possible to visualise

mEos3.1 tagged Rrp1 but it would be very interesting to see if it was behaving in the same way as Rrp2.

A number of mEos3.1 tagged strains produced by other members of the Carr lab have shown characteristics that suggest mEos3.1 is not as well tolerated as GFP. Strains tagged with mEos3.1 have become diploid during the transformation event (allowing them to retain an untagged copy of the gene), have shown cleavage of the majority of mEos3.1 or have lost the tag entirely. No increased sensitivity was observed in the spot tests with *rrp1*-mEos3.1 but, if the cells were either cleaving mEos3.1 from the Rrp1 protein and degrading the free mEos3.1 or somehow recombining the mEos3.1 sequence out of the genome then any sensitivity would have been removed. It is possible that a different fluorophore would be less problematic. There is a photo-activatable GFP available which should be structurally very similar to GFP. A variety of other fluorescent proteins also exist that could be tried, although experiments in the lab show that none have the photon budget of mEos3.1. Alternatively there are a range of organic fluorophores that are much brighter than any of the fluorescent proteins and can be conjugated to antibodies or nanobodies or used in combination with HALO or SNAP tags. However it can be more difficult to introduce organic fluorophores into cells (see Introduction 1.2.3.1).

As it has been assumed that Rrp1 and Rrp2 function as a heterodimer it would be interesting to explore whether Rrp1 would show the same DNA-association patterns under various treatments as was seen with Rrp2. The fact that untreated Δ*rrp1* cells showed a lower number of Rrp2 localisations per nucleus than wild type seems to confirm that the two proteins are linked.

5.3.6.2 Probing other proteins in the Swi5/Sfr1 arm of Rad51-dependent HR

Rrp2 is proposed to function in the Swi5/Sfr1 arm of Rad51-dependent HR and it would be interesting to perform further motion blur experiments in *swi5* and *sfr1* delete backgrounds. No additional sensitivity to DNA damage was seen by Dziadkowiec *et al.*, when *swi5* was deleted in combination with *rrp1* or *rrp2* so it may be that it would not have an effect. However if the Swi5/Sfr1 complex is working together with Rrp2 and Rrp1 to repair damage it might be expected that a change in how the Rrp2 is localised after damage would be observed. If no change in number of Rrp2 localisations were observed it would suggest that Rrp2 is functioning upstream of the Swi5/Sfr1 complex, alternatively if the number of localisations did change it would suggest that Rrp2 is functioning downstream from Swi5/Sfr1.

Additionally, it would be interesting to tag Swi5 or Sfr1 with mEos and perform motion blur on those strains to assess whether their DNA-association followed a similar pattern to Rrp2. Based on the observations of Akamatsu *et al.*, (2007) who observed Swi5 and Sfr1 to form foci after DNA damage it might be expected that the behaviour of Swi5 and Sfr1 will be more like that of Rad52.

5.3.6.3 Super-resolution two-colour co-localisation

Dziadkowiec *et al.,* suggested that Rrp1 and Rrp2 were being localised to sites of damage by using two-colour co-localisation with Rad51, which has been shown to localise to sites of DNA damage. Recently, a few groups have used super-resolution imaging to improve co-localisation imaging (for example Georgieva *et al.*, 2016). Where resolution is diffraction limited it is only possible to say that two proteins are co-localised with an uncertainty of 200 nm or more, which is quite a large distance on a cellular scale, whereas PALM imaging

can give a resolution down to 20-30 nm meaning it is possible to say more strongly whether two things are co-localised. Doing this requires that the two proteins to be observed are tagged with two photo-activatable fluorophores with different excitation and emission properties. This is somewhat challenging, particularly when using genetically encoded fluorophores, as the excitation wavelength of one cannot match the activation wavelength of the other but the emission wavelengths have to be sufficiently different to be separable. Assuming two suitable fluorophores are found the beam path has to be split into two before the camera, one path for each colour, and the shift between these two paths needs to be carefully calculated in order to align the final images properly.

The system used in the Carr lab is not currently set up to do this, but it may become possible to do such two-colour imaging in the future. With such a system available it would be interesting to look at the co-localisation of Rrp2 with Rad51 to start with, to see if what was observed would match the observations of Dziadkowiec *et al.* Then, assuming a different fluorescent tag had enabled the visualisation of Rrp1, it would be interesting to see whether Rrp1 and Rrp2 co-localise on the 10s of nm scale.

5.3.7 Summary

In summary, it has been shown that by using PALM it is possible to study the protein Rrp2 at endogenous levels, which was not possible with conventional fluorescence microscopy. Furthermore, it has been possible to elucidate the behaviour of Rrp2 in response to DNA damage. The analyses show that Rrp2 is constitutively associated with the DNA and that a portion is disassociated in response to DNA damage. This disassociation was not affected by use of different DNA damaging agents or by deficiencies in checkpoint activation. The initial association with chromatin observed in undamaged cells was not affected by heterochromatin formation. The full function of Rrp2 remains to be elucidated and very

recent unpublished data from other groups suggests it may have a different function than was first supposed. Motion blur PALM has proven to be a useful tool that could be put to further use in future experiments.

6 Discussion

6.1 Overview

The main aim of this thesis was to establish how single molecule localisation microscopy (SMLM), more specifically photo-activated localisation microscopy (PALM), could be applied to different types of biological problem in the model organism *S. pombe*.

When this work was initiated a number of papers had been published that utilised PALM to quantify proteins found on membranes or in small numbers in foci. It was hypothesised that it would be possible to use this technique to deduce the stoichiometry of protein complexes in the cytoplasm. A combination of modelling and imaging of artificial oligomers of the fluorescent protein mEos3.1 suggest that it is indeed possible to count proteins in the cytoplasm, but only under very limited circumstances. The density of the protein of interest must be very low and the activation efficiency of the fluorophore used should ideally be higher than mEos3.1. This ruled quantitative PALM out for use in the biological question of determining the stoichiometry of $Cdc22^{R1}$, a subunit of the ribonucleotide reductase (RNR) complex, which is expressed very highly throughout the cytoplasm (Carpy *et al.*, 2014).

In order to further consider the problem of $Cdc22^{R1}$ complex stoichiometry a variation of the technique called fluorescence correlation spectroscopy (FCS) was explored, which can be used to measure diffusion coefficients (from which size and thus stoichiometry can be extrapolated) of fluorescently tagged proteins *in vivo*. Using model substrates evidence has been presented here that it is possible to tell the difference between the diffusion speeds of molecules with only a small difference in size. However, the experiments also suggested that it is harder to separate out two similarly sized molecules in a mixture. When exploring $Cdc22^{R1}$ tagged with mEos3.2 the values seen from analysing the FCS data, combined with

western blot data, suggested that the majority of the mEos3.2 tag was cleaved from $Cdc22^{R1}$ reducing the usefulness of this particular tag-technique combination.

Returning to PALM a different biological problem was proposed to test the limitations of the technique for use in *S. pombe*. A collaborator's group had discovered two proteins, Rrp1 and Rrp2, thought to be involved in homologous recombination (HR), which they could not observe with conventional fluorescence microscopy at endogenous levels. These proteins, when over-expressed from a plasmid, had been observed to form foci (clusters of multiple molecules) at sites of DNA damage. Using PALM it was possible to observe the Rrp2 protein at endogenous levels in fixed cells, although foci formation was not observed. Further, using motion blur, a technique established by Dr Thomas Etheridge (University of Sussex), it was possible to show that Rrp2 is associated with the DNA under normal circumstances and a portion is disassociated as a result of DNA damage.

In conclusion this thesis has demonstrated three things: That there are definite limitations to the quantitative use of PALM techniques inside cells. That FCS, when combined with photo-activation, can be used to distinguish between similarly sized populations *in vivo*, although it may not be so useful where the population is mixed. That PALM is useful for studying proteins with such low expression levels they are hard to see under conventional microscopy.

6.2 Quantitative PALM for cytoplasmic proteins in *S. pombe*

Prior to the development of super-resolution microscopy techniques the diffraction limit of conventional fluorescence microscopy meant that quantitative microscopy could only be performed *in vivo* in bulk, e.g. by measurements of overall fluorescence intensity in the cell

when compared to a standard (reviewed in Coffman and Wu, 2012) or by step-wise photo-bleaching (Ulbrich and Isacoff, 2007). Measuring fluorescence intensity can give a measure of cellular content but not of complex stoichiometry, step-wise photo-bleaching can give information about complex stoichiometry but is limited to 4-5 subunits. Alternative approaches *in vitro* such as cryo-electron microscopy (EM) give more detailed results but as the structure of complexes in cells is often strongly dependent on the presence of scaffold proteins, and other specific conditions that are hard to replicate *in vitro*, results cannot be relied upon to exactly replicate what happens *in vivo*. SMLM super-resolution techniques rely on imaging each fluorescently tagged molecule separately over time, as such it should be theoretically possible to count the number of molecules that have the same (or very close) localisation positions. This has been done using membrane localised proteins (Nan *et al.*, 2013; Puchner *et al.*, 2013; Renz *et al.*, 2012) and studying proteins which form a focus (Lando *et al.*, 2012).

However, the precision which can be achieved studying membrane proteins is a lot higher than for studying cytoplasmic or nuclear proteins. This is because total internal reflection fluorescence (TIRF) microscopy, which can only be used to focus on the surface of a cell at the interface with the glass slide, can be used. This reduces background fluorescence, improving the signal to noise ratio (SNR) and therefore improving precision. Similarly, where proteins exist in a single focus, a significant amount of background can be excluded by simply filtering out everything not in the focus area. Additionally it is not complicated by overlapping of two separate clusters of molecules. Quantitative PALM was considered to be potentially useful to study a protein called $Cdc22^{R1}$, a subunit of the RNR complex, which is highly expressed throughout the cytoplasm and may exist in more than one stoichiometry, with stoichiometry possibly having a role in activity regulation. To initially test the limitations of the technique genetic constructs coding for a monomer, a dimer and a trimer of the photo-activatable fluorescent protein (PA-FP) mEos3.1 were produced and

expressed in *S. pombe* cells. Using plugins written by Dr Alex Herbert (University of Sussex) that worked with the software FIJI (Schindelin *et al.*, 2012) it was found that it was not possible to identify dimers or trimers with a continuous activation scheme. Using a variety of tracing parameters the correct cluster size (*N*) was rarely returned and the value for *p* (photo-activation efficiency) was consistently much lower than expected. Using a pulsed activation scheme, to reduce repeat localisations from blinking, combined with stricter tracing, it was found that under-counting of *N* occurred. This was possibly a result of fitting monomers of noise, cleavage of the multimers or the low photo-activation efficiency of mEos3.1 (Durisic *et al.*, 2014). At the same time modelling performed collaboratively with Dr Herbert demonstrated that high densities of complexes would be impossible to accurately count due to overlapping of separate complexes.

Quantification of cytoplasmic proteins might be possible given a better fluorophore. Durisic *et al.*, (2014) quantified the photo-activation efficiency for a variety of PA-FPs. The PA-FP mEos2 gave a higher photo-activation efficiency than mEos3.1/2, with a score of 60% which could make it a better choice for quantitative PALM. However, when the trimer construct was re-made using mEos2 and imaged using pulsed activation, the PC-PALM clustering algorithm under-counted *N* and gave low *p* values similar to the results seen with mEos3.1. The PA-FPs mMaple (McEvoy *et al.*, 2012) and Dendra2 (Gurskaya *et al.*, 2006) both scored more highly on photo-activation efficiency than mEos2 or 3, it is possible that with an even higher photo-activation efficiency counting might be more effective. The major problem with changing PA-FP is that the Eos derived proteins are among the brightest of the currently available PA-FPs and the higher the brightness, or photon budget, the better the precision of localisation. Dr Herbert's modelling showed that precision was the largest factor in determining whether cluster sizes could be accurately determined.

6.3 Exploring FCS to determine the stoichiometry of the RNR complex

Since it was not possible to determine the stoichiometry of $Cdc22^{R1}$ in complex using PALM, a technique called FCS, which also uses fluorescent tags to study proteins and can be used to gain a measure of the diffusion coefficient which correlates with size, was implemented. It was hypothesised that two or more different diffusion speeds would be observed indicating the different stoichiometries of the RNR complex. FCS had been performed by a number of different groups to study changes in size brought about by protein binding (e.g. Zhong *et al.*, 2001) and protein oligomerisation (e.g. Patel *et al.*, 2002) *in vivo* (reviewed by Mütze *et al.*, 2011). Additionally, because FCS relies on a low number of molecules being measured at any given time, combining FCS measurements with use of PA-FPs had been shown to be beneficial (Kaur *et al.*, 2013b).

FCS works by using the fluctuations in fluorescence in a diffraction limited laser volume as a measure of how molecules are diffusing or changing their fluorescent state within the volume. Fluctuations are recorded and then an autocorrelation function performed on them. The autocorrelation function produces a decay curve from which it is possible to extract information such as the time fluorescent molecules spend in the laser volume. To extract these values it is necessary to fit the decay curve with a model of the system. Conventionally this has been done using an equation that makes an assumption about the number of components with different diffusion times and least squares fitting. However, this approach requires assumptions to be made about the number of components and is limited to approximately four different components, above which it fails. A method of analysis that used a maximum entropy fitting algorithm, which makes no assumptions

about the number of components and has been shown to be useful for analysing highly heterologous populations (Sengupta *et al.*, 2003), was also used. A piece of software was obtained that provided tools to fit the data with both a conventional fit and a maximum entropy fit (Jan Wolfgang Krieger and Jorg Langowski, 2015).

The protocol was benchmarked using the monomer, dimer and trimer of mEos3.1 produced for the PALM counting experiments and I was able to show that differentiation was possible between things of very similar sizes *in vivo* using maximum entropy fitting and also, to a lesser extent, using conventional fitting. Since diffusion coefficient scales with roughly the cube root of the molecular weight (MW) (Lakowicz, 2006) it is generally assumed that the difference in molecular weight needs to be large to see a difference in diffusion. As such the expected difference in diffusion time between a monomer and a trimer would be on the scale of $3^{1/3}$, in other words a 1.44 fold difference, and less than that between monomer and dimer. Previous studies on the resolution of FCS have suggested that a difference in diffusion speed of at least 1.6 fold is required (Meseth *et al.*, 1999). However, a difference was observed between monomer and trimer diffusion times using conventional fitting and a difference between monomer, dimer and trimer using the maximum entropy fitting method.

Having demonstrated that it was possible to differentiate between species that had only a small difference in MW when measured separately it was necessary to ask whether it would still be possible when the populations were mixed. Results with rhodamine B labelled dextrans of 3 kDa and 10 kDa, measured alone and in mixtures of different proportions, suggested that this aspect would be more difficult. A future experiment it would be good to perform would be to express the monomer and the trimer constructs from the same cell line. However, this would need to be carefully controlled as it is possible that inserting both constructs into the genome could lead to recombination between them (due to sequence

similarity) which would change the sequence of the construct.

Finally looking at Cdc22^{R1} tagged with mEos3.2 posed a problem as it was found that the Cdc22^{R1} cells were able to reject the tag and that, in cells in which the tag was present, a large percentage of it was cleaved from Cdc22^{R1}. This was borne out in the FCS results as both conventional fitting and maximum entropy fitting returned a value consistent with the diffusion speed seen for a monomer of mEos3.1. However, maximum entropy fitting did provide a third peak at a very long time which may have represented a form of the RNR complex and this third long time peak was not seen in cells which had been HU treated indicating it may indeed be reflecting RNR. One possible problem is that if this third peak does represent the RNR complex it may be at diffusion times where the limit of capacity to measure is reached because the fluorophore bleaches before it crosses the volume.

The problem of mEos3.2 cleavage and rejection may mean that it is not a suitable tag for using with FCS to study Cdc22^{R1}. There have been various other *S. pombe* strains produced within the Carr research group where a protein that has appeared to be functional with a GFP tag has had its function compromised by mEos, suggesting that it is the tag specifically that is the problem. It would be worthwhile to try a different PA-FP, possibly PA-GFP to see if that was cleaved less, or would not be rejected as readily. If a more stable tag could be found there are a number of experiments that might be done to look for a change in the maximum entropy profile. It would be particularly interesting to look at the effect of deleting the gene for the small protein Spd1, which has several roles in regulating RNR activity, has been shown to bind to Cdc22^{R1} (Hakansson *et al.*, 2006) and was observed to be necessary for förster resonance energy transfer (FRET) between Cdc22^{R1} and the other subunit of the RNR complex, Suc22^{R2} (Nestoras *et al.*, 2010). As RNR activity must be regulated in time with the cell cycle it would also be interesting to use lactose gradient synchronisation to enrich for populations in G2 or S phase, take FCS

measurements in a number of cells at each stage and compare them using the maximum entropy fitting.

Although this project did not result in concrete answers about the stoichiometry of the RNR complex it did provide evidence that FCS can be used to distinguish things more similar in size than initially anticipated. With a better fluorescent tag there is the potential to elucidate more information about Cdc22^{R1} and the RNR complex. Hopefully these results will encourage other people to make use of this technique for similar biological questions.

6.4 Using PALM to study the low abundance protein Rrp2 at endogenous levels

The fission yeast *rrp1* and *rrp2* genes were discovered by Dziadkowiec *et al.*, (2009) in a search for sequence homology to budding yeast *ris1*, a DNA-dependent ATPase with roles in DNA repair and gene silencing. They showed that the proteins Rrp1 and Rrp2 are involved in the Swi5/Sfr1 arm of homologous recombination and that, when over-expressed from a plasmid, they appear to form foci (clusters of multiple molecules) at sites of damage (Dziadkowiec *et al.*, 2013). At endogenous expression levels it was found to be very hard to visualise Rrp2 using conventional fluorescence microscopy. In this thesis it has been established, using PALM and the motion blur technique (Etheridge *et al.*, 2014), that no obvious Rrp2 foci can be seen in fixed cells that have been treated with methane methylsulphonate (MMS), that Rrp2 is associated with the DNA under normal conditions and that less Rrp2 is associated with the DNA after treatment with MMS, HU and ionising radiation. These observations were not found to change as a result of deletion of genes relating to checkpoint pathway activation or heterochromatin assembly. However, it seemed that less Rrp2 was associated in the undamaged state in the absence of either Rad57 (the

alternative arm of the Rad51-dependent homologous recombination pathway) or Rrp1. Whilst this does not clarify the regulation or function of Rrp2 it does reduce the possibilities; it is not being disassociated from DNA by checkpoint genes and it does not require heterochromatin formation for association or disassociation. When it is involved in HR less of it is required than is associated with the DNA in the absence of damage.

Dziadkowiec *et al.*, used a yeast two-hybrid screen and co-localisation microscopy to suggest that Rrp1 and Rrp2 work together. It would be of great interest to know whether Rrp1 behaves in the same way as Rrp2 in terms of DNA-association. Although a strain expressing mEos3.1 tagged Rrp1 was produced it was not possible to detect Rrp1 using the PALM microscope. The possibility was considered that there might simply be much less of Rrp1 to the extent even super-resolution microscopy cannot help to visualise it. According to Pombase.org the number of RNA molecules per cell during vegetative growth is 2.3 for Rrp1 and 1.1 for Rrp2 (Marguerat *et al.*, 2012). The number of localisations for Rrp2 had an average of ~100 in unperturbed cells, this is likely to be an under-estimate of absolute numbers due to the photo-activation efficiency of mEos. That this is 100 fold higher than the reported transcript level suggests the protein is reasonably stable. If it is assumed the quantity of transcript correlates with the quantity of protein, it might be expected that there would be twice as much Rrp1 as Rrp2 but this is not what was observed. The Rrp1 protein was tagged at the N-terminus whereas Rrp2 was tagged at the C-terminus, it is possible that this has had an effect on the ability of mEos3.1 to fold correctly leading to non-fluorescent mEos3.1.

Collaborators in the Dziadkowiec group are currently working on creating a series of point mutations in *rrp1* and *rrp2* in areas whose sequences suggest functional importance such as the ATPase domain. Once these have been produced they could be cloned into the cassette exchange plasmid, tagged with mEos and cassette exchanged into the base strain in order

to see how the mutations affect the DNA-association patterns that have been observed.

As it has been demonstrated that Rrp2 functions in the Swi5/Sfr1 pathway during HR it would be worthwhile to examine the Rrp2 DNA-association in strains in which either *swi5*, *sfr1* or both had been deleted. Additionally, Swi5 and Sfr1 have also been shown to form foci that co-localise with Rad51 foci with Swi5 forming some foci spontaneously (Akamatsu *et al.*, 2007). It would be very interesting to tag Swi5 with mEos3.1 and image it using the motion blur protocol with the same treatments and backgrounds used with Rrp2 to establish whether it shows similar association patterns.

Whilst this thesis was being written unpublished data from the research group of Dr Li Lin Du identified a new role for Rrp2, the removal of SUMO-modified topoisomerase Top2 after CPT treatment. This function does not require Rrp1 and potentially provides an explanation for the observations made here concerning the DNA or chromatin association of Rrp2 in undamaged cells. It would be interesting to explore the requirements for the SUMO modification pathway and for Top2 itself in this DNA-association and to explore the effects of treating cells with topoisomerase poison.

This project has established that PALM can be useful for studying proteins with very low endogenous expression levels. This is important because over-expression from a plasmid risks upsetting the internal balance of the cell as has been seen by the Dziadkowiec group who have now observed that long term over-expression of Rrp1 can lead to growth defects. Additionally binding kinetics are dependent on concentration and as such the associations observed in an over-expressed protein may only be vaguely related to what occurs at endogenous levels. Where possible it is always most biologically informative to observe things *in vivo* and at endogenous levels.

7. Bibliography

1951 USAF resolution test chart, 2016. . Wikipedia Free Encycl.

Abbe, E., 1873. Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. Arch. Für Mikrosk. Anat. 9, 413–418. doi:10.1007/BF02956173

Abraham, A.V., Ram, S., Chao, J., Ward, E.S., Ober, R.J., 2009. Quantitative study of single molecule location estimation techniques. Opt. Express 17, 23352–23373.

Adam, V., 2014. Phototransformable fluorescent proteins: which one for which application? Histochem. Cell Biol. 142, 19–41. doi:10.1007/s00418-014-1190-5

Adam, V., Lelimousin, M., Boehme, S., Desfonds, G., Nienhaus, K., Field, M.J., Wiedenmann, J., McSweeney, S., Nienhaus, G.U., Bourgeois, D., 2008. Structural characterization of IrisFP, an optical highlighter undergoing multiple photo-induced transformations. Proc. Natl. Acad. Sci. 105, 18343–18348. doi:10.1073/pnas.0805949105

Adam, V., Nienhaus, K., Bourgeois, D., Nienhaus, G.U., 2009. Structural Basis of Enhanced Photoconversion Yield in Green Fluorescent Protein-like Protein Dendra2. Biochemistry (Mosc.) 48, 4905–4915. doi:10.1021/bi900383a

Airy (Sir), G.B., 1835. On the Diffraction of an Object-glass with Circular Aperture. printed at the Pitt Press, by J. Smith.

Akamatsu, Y., Dziadkowiec, D., Ikeguchi, M., Shinagawa, H., Iwasaki, H., 2003. Two different Swi5-containing protein complexes are involved in mating-type switching and recombination repair in fission yeast. Proc. Natl. Acad. Sci. 100, 15770–15775.

Akamatsu, Y., Tsutsui, Y., Morishita, T., Siddique, M.S.P., Kurokawa, Y., Ikeguchi, M., Yamao, F., Arcangioli, B., Iwasaki, H., 2007. Fission yeast Swi5/Sfr1 and Rhp55/Rhp57 differentially regulate Rhp51-dependent recombination outcomes. EMBO J. 26, 1352–1362. doi:10.1038/sj.emboj.7601582

Amos, W. b., White, J. g., 2003. How the Confocal Laser Scanning Microscope entered Biological Research. Biol. Cell 95, 335–342. doi:10.1016/S0248-4900(03)00078-9

Annibale, P., Scarselli, M., Kodiyan, A., Radenovic, A., 2010. Photoactivatable Fluorescent Protein mEos2 Displays Repeated Photoactivation after a Long-Lived Dark State in the Red Photoconverted Form. J. Phys. Chem. Lett. 1, 1506–1510. doi:10.1021/jz1003523

Annibale, P., Vanni, S., Scarselli, M., Rothlisberger, U., Radenovic, A., 2011. Quantitative Photo Activated Localization Microscopy: Unraveling the Effects of Photoblinking. PLoS ONE 6, e22678. doi:10.1371/journal.pone.0022678

Aragón, S.R., Pecora, R., 1975. Fluorescence correlation spectroscopy and Brownian rotational diffusion. Biopolymers 14, 119–137. doi:10.1002/bip.1975.360140110

Arita, Y., Nishimura, S., Matsuyama, A., Yashiroda, Y., Usui, T., Boone, C., Yoshida, M., 2011. Microarray-based target identification using drug hypersensitive fission yeast expressing ORFeome. Mol. Biosyst. 7, 1463. doi:10.1039/c0mb00326c

Arnaudeau, C., Lundin, C., Helleday, T., 2001. DNA double-strand breaks associated with replication forks are predominantly repaired by homologous recombination involving an exchange mechanism in mammalian cells1. J. Mol. Biol. 307, 1235–1245. doi:10.1006/jmbi.2001.4564

Arnaudeau, C., Tenorio Miranda, E., Jenssen, D., Helleday, T., 2000. Inhibition of DNA synthesis is a potent mechanism by which cytostatic drugs induce homologous recombination in mammalian cells. Mutat. Res. 461, 221–228.

Axelrod, D., 1981. Cell-substrate contacts illuminated by total internal reflection fluorescence. J. Cell Biol. 89, 141–145.

Baddeley, D., Jayasinghe, I.D., Lam, L., Rossberger, S., Cannell, M.B., Soeller, C., 2009. Optical single-channel resolution imaging of the ryanodine receptor distribution in rat cardiac myocytes. Proc. Natl. Acad. Sci. U. S. A. 106, 22275–22280. doi:10.1073/pnas.0908971106

Baeyer, A., 1871. Ueber eine neue Klasse von Farbstoffen. Berichte Dtsch. Chem. Ges. 4, 555–558. doi:10.1002/cber.18710040209

Bähler, J., Wu, J.Q., Longtine, M.S., Shah, N.G., McKenzie, A., Steever, A.B., Wach, A., Philippsen, P., Pringle, J.R., 1998. Heterologous modules for efficient and versatile PCR-based gene targeting in Schizosaccharomyces pombe. Yeast Chichester Engl. 14, 943–951. doi:10.1002/(SICI)1097-0061(199807)14:10<943::AID-YEA292>3.0.CO;2-Y

Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, R.C., Kouzarides, T., 2001. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature 410, 120–124.

Behbehani, G.K., Thom, C., Zunder, E.R., Finck, R., Gaudilliere, B., Fragiadakis, G.K., Fantl, W.J., Nolan, G.P., 2014. Transient partial permeabilization with saponin enables cellular barcoding prior to surface marker staining. Cytom. Part J. Int. Soc. Anal. Cytol. 85, 1011–1019. doi:10.1002/cyto.a.22573

Bentley, N.J., Holtzman, D.A., Flaggs, G., Keegan, K.S., DeMaggio, A., Ford, J.C., Hoekstra, M., Carr, A.M., 1996. The Schizosaccharomyces pombe rad3 checkpoint gene. EMBO J. 15, 6641.

Berg, J.M., Tymoczko, J.L., Stryer, L., 2006. Biochemistry 6th Edition, 6th ed.

Berland, K.M., So, P.T., Gratton, E., 1995. Two-photon fluorescence correlation spectroscopy: method and application to the intracellular environment. Biophys. J. 68, 694.

Betzig, E., Patterson, G.H., Sougrat, R., Lindwasser, O.W., Olenych, S., Bonifacino, J.S., Davidson, M.W., Lippincott-Schwartz, J., Hess, H.F., 2006a. Imaging Intracellular Fluorescent Proteins at Nanometer Resolution. Science 313, 1642–1645. doi:10.1126/science.1127344

Betzig, E., Patterson, G.H., Sougrat, R., Lindwasser, O.W., Olenych, S., Bonifacino, J.S., Davidson, M.W., Lippincott-Schwartz, J., Hess, H.F., 2006b. Imaging Intracellular Fluorescent Proteins at Nanometer Resolution. Science 313, 1642–1645. doi:10.1126/science.1127344

Bielawski, K., Wołczyński, S., Bielawska, A., 2001. DNA-binding activity and cytotoxicity of the extended diphenylfuran bisamidines in breast cancer MCF-7 cells. Biol. Pharm. Bull. 24, 704–706.

Bisht, K.K., Arora, S., Ahmed, S., Singh, J., 2008. Role of heterochromatin in suppressing subtelomeric recombination in fission yeast. Yeast 25, 537–548. doi:10.1002/yea.1603

Biteen, J.S., Moerner, W.E., 2010. Single-Molecule and Superresolution Imaging in Live Bacteria Cells. Cold Spring Harb. Perspect. Biol. 2, a000448. doi:10.1101/cshperspect.a000448

Biteen, J.S., Thompson, M.A., Tselentis, N.K., Bowman, G.R., Shapiro, L., Moerner, W.E., 2008. Super-resolution imaging in live Caulobacter crescentus cells using photoswitchable EYFP. Nat. Methods 5, 947–949. doi:10.1038/nmeth.1258

Björklund, S., Skog, S., Tribukait, B., Thelander, L., 1990. S-phase-specific expression of mammalian ribonucleotide reductase R1 and R2 subunit mRNAs. Biochemistry (Mosc.) 29, 5452–5458.

Bobroff, N., 1986. Position measurement with a resolution and noise-limited instrument. Rev. Sci. Instrum. 57, 1152. doi:10.1063/1.1138619

Boddy, M.N., Furnari, B., Mondesert, O., Russell, P., 1998. Replication checkpoint enforced by kinases Cds1 and Chk1. Science 280, 909–912.

Brancaleon, L., Durkin, A.J., Tu, J.H., Menaker, G., Fallon, J.D., Kollias, N., 2001. In vivo fluorescence spectroscopy of nonmelanoma skin cancer. Photochem. Photobiol. 73, 178–183.

Branzei, D., Foiani, M., 2010. Maintaining genome stability at the replication fork. Nat. Rev. Mol. Cell Biol. 11, 208–219. doi:10.1038/nrm2852

Caldecott, K.W., 2008. Single-strand break repair and genetic disease. Nat. Rev. Genet. 9, 619–631. doi:10.1038/nrg2380

Cann, K.L., Dellaire, G., 2010. Heterochromatin and the DNA damage response: the need to relax. Biochem. Cell Biol. 89, 45–60. doi:10.1139/O10-113

Carpy, A., Krug, K., Graf, S., Koch, A., Popic, S., Hauf, S., Macek, B., 2014. Absolute Proteome and Phosphoproteome Dynamics during the Cell Cycle of Schizosaccharomyces pombe (Fission Yeast). Mol. Cell. Proteomics 13, 1925–1936. doi:10.1074/mcp.M113.035824

Caspari, T., Carr, A.M., 1999. DNA structure checkpoint pathways in Schizosaccharomyces pombe. Biochimie 81, 173–181.

Chabes, A., Domkin, V., Thelander, L., 1999. Yeast Sml1, a protein inhibitor of ribonucleotide reductase. J. Biol. Chem. 274, 36679–36683.

Chen, B.-C., Legant, W.R., Wang, K., Shao, L., Milkie, D.E., Davidson, M.W., Janetopoulos, C., Wu, X.S., Hammer, J.A., Liu, Z., English, B.P., Mimori-Kiyosue, Y., Romero, D.P., Ritter, A.T., Lippincott-Schwartz, J., Fritz-Laylin, L., Mullins, R.D., Mitchell, D.M., Bembenek, J.N., Reymann, A.-C., Böhme, R., Grill, S.W., Wang, J.T., Seydoux, G., Tulu, U.S., Kiehart, D.P., Betzig, E., 2014. Lattice Light Sheet Microscopy: Imaging Molecules to Embryos at High Spatiotemporal Resolution. Science 346, 1257998. doi:10.1126/science.1257998

Chen, H., Farkas, E.R., Webb, W.W., 2008. Chapter 1 In Vivo Applications of Fluorescence Correlation Spectroscopy, in: Methods in Cell Biology. Elsevier, pp. 3–35.

Chen, Y., Müller, J.D., So, P.T., Gratton, E., 1999. The photon counting histogram in fluorescence fluctuation spectroscopy. Biophys. J. 77, 553–567.

Chen, Y., Wei, L.-N., Müller, J.D., 2003. Probing protein oligomerization in living cells with fluorescence fluctuation spectroscopy. Proc. Natl. Acad. Sci. U. S. A. 100, 15492–15497. doi:10.1073/pnas.2533045100

Cheutin, T., McNairn, A.J., Jenuwein, T., Gilbert, D.M., Singh, P.B., Misteli, T., 2003. Maintenance of Stable Heterochromatin Domains by Dynamic HP1 Binding. Science 299, 721–725. doi:10.1126/science.1078572

Christensen, P.U., Bentley, N.J., Martinho, R.G., Nielsen, O., Carr, A.M., 2000. Mik1 levels accumulate in S phase and may mediate an intrinsic link between S phase and mitosis. Proc. Natl. Acad. Sci. U. S. A. 97, 2579–2584.

Coffman, V.C., Wu, J.-Q., 2012. Counting protein molecules using quantitative fluorescence microscopy. Trends Biochem. Sci. 37, 499–506. doi:10.1016/j.tibs.2012.08.002

Compare 2 Means 2-Sample, 2-Sided Equality | Power and Sample Size Calculators | HyLown [WWW Document], 2016. URL http://powerandsamplesize.com/Calculators/Compare-2-Means/2-Sample-Equality (accessed 7.13.16).

Coons, A., Creech, H., Jones, R.N., Berliner, E., 1942. The Demonstration of Pneumococcal Antigen in Tissues by the Use of Fluorescent Antibody. J. Immunol. 45, 159–170.

Coons, A.H., Creech, H.J., Jones, R.N., 1941. Immunological Properties of an Antibody Containing a Fluorescent Group. Exp. Biol. Med. 47, 200–202. doi:10.3181/00379727-47-13084P

Costanzo, V., 2011. Brca2, Rad51 and Mre11: Performing balancing acts on replication forks. DNA Repair 10, 1060–1065. doi:10.1016/j.dnarep.2011.07.009

Cremer, T., Cremer, C., 2001. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. Nat. Rev. Genet. 2, 292–301. doi:10.1038/35066075

Croft, W.J., 2006. Under the Microscope: A Brief History of Microscopy. World Scientific.

Dalgaard, J.Z., Klar, A.J.S., 2001. A DNA replication-arrest site RTS1 regulates imprinting by determining the direction of replication at mat1 in S. pombe. Genes Dev. 15, 2060–2068. doi:10.1101/gad.200801

Delacroix, S., Wagner, J.M., Kobayashi, M., Yamamoto, K., Karnitz, L.M., 2007. The Rad9–Hus1–Rad1 (9–1–1) clamp activates checkpoint signaling via TopBP1. Genes Dev. 21, 1472–1477. doi:10.1101/gad.1547007

Denk, W., Strickler, J.H., Webb, W.W., others, 1990. Two-photon laser scanning fluorescence microscopy. Science 248, 73–76.

Dou, H., Huang, C., Van Nguyen, T., Lu, L.-S., Yeh, E.T.H., 2011. SUMOylation and de-SUMOylation in response to DNA damage. FEBS Lett. 585, 2891–2896. doi:10.1016/j.febslet.2011.04.002

Durisic, N., Laparra-Cuervo, L., Sandoval-Álvarez, Á., Borbely, J.S., Lakadamyali, M., 2014. Single-molecule evaluation of fluorescent protein photoactivation efficiency using an in vivo nanotemplate. Nat. Methods 11, 156–162. doi:10.1038/nmeth.2784

Dziadkowiec, D., Kramarz, K., Kanik, K., Wisniewski, P., Carr, A.M., 2013. Involvement of Schizosaccharomyces pombe rrp1+ and rrp2+ in the Srs2- and Swi5/Sfr1-dependent pathway in response to DNA damage and replication inhibition. Nucleic Acids Res. 41, 8196–8209. doi:10.1093/nar/gkt564

Dziadkowiec, D., Petters, E., Dyjankiewicz, A., Karpiński, P., Garcia, V., Watson, A., Carr, A.M., 2009. The role of novel genes rrp1+ and rrp2+ in the repair of DNA damage in Schizosaccharomyces pombe. DNA Repair 8, 627–636. doi:10.1016/j.dnarep.2008.12.008

Edelstein, A.D., Tsuchida, M.A., Amodaj, N., Pinkard, H., Vale, R.D., Stuurman, N., 2014. Advanced methods of microscope control using μManager software. J. Biol. Methods 1, 10. doi:10.14440/jbm.2014.36

Elgin, S.C., Grewal, S.I., 2003. Heterochromatin: silence is golden. Curr. Biol. 13, R895–R898.

Ellison, V., Stillman, B., 2003. Biochemical Characterization of DNA Damage Checkpoint Complexes: Clamp Loader and Clamp Complexes with Specificity for 5' Recessed DNA. PLoS Biol. 1, e33. doi:10.1371/journal.pbio.0000033

Elowitz, M.B., Surette, M.G., Wolf, P.-E., Stock, J.B., Leibler, S., 1999. Protein Mobility in the Cytoplasm of Escherichia coli. J. Bacteriol. 181, 197–203.

Elson, E.L., Magde, D., 1974. Fluorescence correlation spectroscopy. I. Conceptual basis and theory. Biopolymers 13, 1–27. doi:10.1002/bip.1974.360130102

Eng, J., Lynch, R.M., Balaban, R.S., 1989. Nicotinamide adenine dinucleotide fluorescence spectroscopy and imaging of isolated cardiac myocytes. Biophys. J. 55, 621–630.

Erami, Z., Timpson, P., Yao, W., Zaidel-Bar, R., Anderson, K.I., 2015. There are four dynamically and functionally distinct populations of E-cadherin in cell junctions. Biol. Open 4, 1481–1489. doi:10.1242/bio.014159

Eriksson, S., Thelander, L., Akerman, M., 1979. Allosteric regulation of calf thymus ribonucleoside diphosphate reductase. Biochemistry (Mosc.) 18, 2948–2952.

Etheridge, T.J., 2015. APPLICATION OF PHOTOACTIVATED LOCALISATION MICROSCOPY TO VISUALISING EUKARYOTIC DNA REPLICATION PROCESSES. University of Sussex.

Etheridge, T.J., Boulineau, R.L., Herbert, A., Watson, A.T., Daigaku, Y., Tucker, J., George, S., Jonsson, P., Palayret, M., Lando, D., Laue, E., Osborne, M.A., Klenerman, D., Lee, S.F., Carr, A.M., 2014. Quantification of DNA-associated proteins inside eukaryotic cells using single-molecule localization microscopy. Nucleic Acids Res. 42, e146–e146. doi:10.1093/nar/gku726

Fairman, J.W., Wijerathna, S.R., Ahmad, M.F., Xu, H., Nakano, R., Jha, S., Prendergast, J., Welin, R.M., Flodin, S., Roos, A., Nordlund, P., Li, Z., Walz, T., Dealwis, C.G., 2011. Structural basis for allosteric regulation of human ribonucleotide reductase by nucleotide-induced oligomerization. Nat. Struct. Mol. Biol. 18, 316–322. doi:10.1038/nsmb.2007

Fischer, T., Cui, B., Dhakshnamoorthy, J., Zhou, M., Rubin, C., Zofall, M., Veenstra, T.D., Grewal, S.I.S., 2009. Diverse roles of HP1 proteins in heterochromatin assembly and functions in fission yeast. Proc. Natl. Acad. Sci. U. S. A. 106, 8998–9003. doi:10.1073/pnas.0813063106

Forsburg, S.L., 2015. Forsburg Lab [WWW Document]. URL http://www-bcf.usc.edu/~forsburg/main4.html (accessed 2.7.17).

Forsburg, S.L., 1994. Codon usage table for Schizosaccharomyces pombe. Yeast 10, 1045–1047. doi:10.1002/yea.320100806

Fuchs, J., Böhme, S., Oswald, F., Hedde, P.N., Krause, M., Wiedenmann, J., Nienhaus, G.U., 2010. A photoactivatable marker protein for pulse-chase imaging with superresolution. Nat. Methods 7, 627–630. doi:10.1038/nmeth.1477

Furnari, B., Blasina, A., Boddy, M.N., McGowan, C.H., Russell, P., 1999. Cdc25 inhibited in vivo and in vitro by checkpoint kinases Cds1 and Chk1. Mol. Biol. Cell 10, 833–845.

Furuya, K., Carr, A.M., 2003. DNA checkpoints in fission yeast. J. Cell Sci. 116, 3847–3848. doi:10.1242/jcs.00790

Gautier, J., Solomon, M.J., Booher, R.N., Bazan, J.F., Kirschner, M.W., 1991. cdc25 is a specific tyrosine phosphatase that directly activates p34cdc2. Cell 67, 197–211. doi:10.1016/0092-8674(91)90583-K

Georgieva, M., Cattoni, D.I., Fiche, J.-B., Mutin, T., Chamousset, D., Nollmann, M., 2016. Nanometer resolved single-molecule colocalization of nuclear factors by two-color super resolution microscopy imaging. Methods San Diego Calif 105, 44–55. doi:10.1016/j.ymeth.2016.03.029

Gerald, J.N.F., Benjamin, J.M., Kron, S.J., 2002. Robust G1 checkpoint arrest in budding yeast: dependence on DNA damage signaling and repair. J. Cell Sci. 115, 1749–1757.

Gest, H., 2007. The Discovery of Microorganisms Revisited.

Gordon, C.B., Fantes, P.A., 1986. The cdc22 gene of Schizosaccharomyces pombe encodes a cell cycle-regulated transcript. EMBO J. 5, 2981.

Gould, K.L., Nurse, P., 1989. Tyrosine phosphorylation of the fission yeast cdc2+ protein kinase regulates entry into mitosis. Nature 342, 39–45. doi:10.1038/342039a0

Gräslund, A., Sahlin, M., Sjöberg, B.M., 1985. The tyrosyl free radical in ribonucleotide reductase. Environ. Health Perspect. 64, 139–149.

Greenfield, D., McEvoy, A.L., Shroff, H., Crooks, G.E., Wingreen, N.S., Betzig, E., Liphardt, J., 2009. Self-Organization of the Escherichia coli Chemotaxis Network Imaged with Super-Resolution Light Microscopy. PLoS Biol. 7. doi:10.1371/journal.pbio.1000137

Grewal, S.I.S., Jia, S., 2007. Heterochromatin revisited. Nat. Rev. Genet. 8, 35–46. doi:10.1038/nrg2008

Gunzenhäuser, J., Olivier, N., Pengo, T., Manley, S., 2012. Quantitative Super-Resolution Imaging Reveals Protein Stoichiometry and Nanoscale Morphology of Assembling HIV-Gag Virions. Nano Lett. 12, 4705–4710. doi:10.1021/nl3021076

Gurskaya, N.G., Verkhusha, V.V., Shcheglov, A.S., Staroverov, D.B., Chepurnykh, T.V., Fradkov, A.F., Lukyanov, S., Lukyanov, K.A., 2006. Engineering of a monomeric green-to-red photoactivatable fluorescent protein induced by blue light. Nat. Biotechnol. 24, 461–465. doi:10.1038/nbt1191

Hakansson, P., Dahl, L., Chilkova, O., Domkin, V., Thelander, L., 2006. The Schizosaccharomyces pombe Replication Inhibitor Spd1 Regulates Ribonucleotide Reductase Activity and dNTPs by Binding to the Large Cdc22 Subunit. J. Biol. Chem. 281, 1778–1783. doi:10.1074/jbc.M511716200

Han, J.J., Kunde, Y.A., Hong-Geller, E., Werner, J.H., 2014. Actin restructuring during Salmonella typhimurium infection investigated by confocal and super-resolution microscopy. J. Biomed. Opt. 19, 016011–016011.

Harigaya, Y., Tanaka, H., Yamanaka, S., Tanaka, K., Watanabe, Y., Tsutsumi, C., Chikashige, Y., Hiraoka, Y., Yamashita, A., Yamamoto, M., 2006. Selective elimination of messenger RNA prevents an incidence of untimely meiosis. Nature 442, 45–50. doi:10.1038/nature04881

Haruta, N., Kurokawa, Y., Murayama, Y., Akamatsu, Y., Unzai, S., Tsutsui, Y., Iwasaki, H., 2006. The Swi5–Sfr1 complex stimulates Rhp51/Rad51 - and Dmc1-mediated DNA strand exchange in vitro. Nat. Struct. 38 Mol. Biol. 13, 823–830. doi:10.1038/nsmb1136

Haustein, E., Schwille, P., 2007. Fluorescence Correlation Spectroscopy: Novel Variations of an Established Technique. Annu. Rev. Biophys. Biomol. Struct. 36, 151–169. doi:10.1146/annurev.biophys.36.040306.132612

Heilemann, M., van de Linde, S., Schüttpelz, M., Kasper, R., Seefeldt, B., Mukherjee, A., Tinnefeld, P., Sauer, M., 2008. Subdiffraction-Resolution Fluorescence Imaging with Conventional Fluorescent Probes. Angew. Chem. Int. Ed. 47, 6172–6176. doi:10.1002/anie.200802376

Hell, S.W., Wichmann, J., 1994. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. Opt. Lett. 19, 780–782.

Herbert, A., 2016. Single-molecule Plugins : ImageJ : ... : Sussex Centre for Genome Damage and Stability : Lifesci : Schools : Staff : University of Sussex [WWW Document]. URL http://www.sussex.ac.uk/gdsc/intranet/microscopy/imagej/smlm_plugins (accessed 6.13.16).

Herschel, J.F.W., 1845. On a Case of Superficial Colour Presented by a Homogeneous Liquid Internally Colourless. Philos. Trans. R. Soc. Lond. 135, 143–145. doi:10.1098/rstl.1845.0004

Hess, S.T., Girirajan, T.P.K., Mason, M.D., 2006. Ultra-High Resolution Imaging by Fluorescence Photoactivation Localization Microscopy. Biophys. J. 91, 4258–4272. doi:10.1529/biophysj.106.091116

Hess, S.T., Gould, T.J., Gudheti, M.V., Maas, S.A., Mills, K.D., Zimmerberg, J., 2007. Dynamic clustered distribution of hemagglutinin resolved at 40 nm in living cell membranes discriminates between raft theories. Proc. Natl. Acad. Sci. 104, 17370–17375.

Höfling, F., Franosch, T., 2013. Anomalous transport in the crowded world of biological cells. Rep. Prog. Phys. 76, 046602. doi:10.1088/0034-4885/76/4/046602

Hofmann, M., Eggeling, C., Jakobs, S., Hell, S.W., 2005. Breaking the diffraction barrier in fluorescence microscopy at low light intensities by using reversibly photoswitchable proteins. Proc. Natl. Acad. Sci. U. S. A. 102, 17565–17569. doi:10.1073/pnas.0506010102

Hooke, R., 1665. Micrographia: or some physiological descriptions of minute bodies made by magnifying glasses : with observations and inquiries thereupon.

Houtgraaf, J.H., Versmissen, J., van der Giessen, W.J., 2006. A concise review of DNA damage checkpoints and repair in mammalian cells. Cardiovasc. Revasc. Med. 7, 165–172. doi:10.1016/j.carrev.2006.02.002

Huisken, J., Swoger, J., Del Bene, F., Wittbrodt, J., Stelzer, E.H., 2004. Optical sectioning deep inside live embryos by selective plane illumination microscopy. Science 305, 1007–1009.

Ingemarson, R., Thelander, L., 1996. A kinetic study on the influence of nucleoside triphosphate effectors on subunit interaction in mouse ribonucleotide reductase. Biochemistry (Mosc.) 35, 8603–8609.

Jamieson, T., Bakhshi, R., Petrova, D., Pocock, R., Imani, M., Seifalian, A.M., 2007. Biological applications of quantum dots. Biomaterials 28, 4717–4732. doi:10.1016/j.biomaterials.2007.07.014

Jamur, M., Oliver, C., 2010. Permeabilization of Cell Membranes, in: Oliver, C., Jamur, M.C. (Eds.), Immunocytochemical Methods and Protocols, Methods in Molecular Biology. Humana Press, pp. 63–66. doi:10.1007/978-1-59745-324-0_9

Jan Wolfgang Krieger, Jorg Langowski, 2015. QuickFit 3.0 (status: beta, compiled: 2015-10-06, SVN: 4373): A data evaluation application for biophysics.

Johnson, I.D., Davidson, M.W., 2012. Olympus Microscopy Resource Center | Jablonski Energy Diagram - Java Tutorial [WWW Document]. URL http://www.olympusmicro.com/primer/java/jablonski/jabintro/ (accessed 8.17.16).

Johnson, R.E., Washington, M.T., Prakash, S., Prakash, L., 1999. Bridging the gap: a family of novel DNA polymerases that replicate faulty DNA. Proc. Natl. Acad. Sci. 96, 12224–12226.

Kai, M., Boddy, M.N., Russell, P., Wang, T.S.-F., 2005. Replication checkpoint kinase Cds1 regulates Mus81 to preserve genome integrity during replication stress. Genes Dev. 19, 919–932. doi:10.1101/gad.1304305

Kara A. Nyberg, Rhett J. Michelson, Charles W. Putnam, Weinert, and T.A., 2002. Toward Maintaining the Genome: DNA Damage and Replication Checkpoints. Annu. Rev. Genet. 36, 617–656. doi:10.1146/annurev.genet.36.060402.113540

Kaur, G., Costa, M.W., Nefzger, C.M., Silva, J., Fierro-González, J.C., Polo, J.M., Bell, T.D.M., Plachta, N., 2013a. Probing transcription factor diffusion dynamics in the living mammalian embryo with photoactivatable fluorescence correlation spectroscopy. Nat. Commun. 4, 1637. doi:10.1038/ncomms2657

Kaur, G., Costa, M.W., Nefzger, C.M., Silva, J., Fierro-González, J.C., Polo, J.M., Bell, T.D.M., Plachta, N., 2013b. Probing transcription factor diffusion dynamics in the living mammalian embryo with photoactivatable fluorescence correlation spectroscopy. Nat. Commun. 4, 1637. doi:10.1038/ncomms2657

Keppler, A., Gendreizig, S., Gronemeyer, T., Pick, H., Vogel, H., Johnsson, K., 2002. A general method for the covalent labeling of fusion proteins with small molecules in vivo. Nat. Biotechnol. 21, 86–89. doi:10.1038/nbt765

Khasanov, F.K., Savchenko, G.V., Bashkirova, E.V., Korolev, V.G., Heyer, W.-D., Bashkirov, V.I., 1999. A New Recombinational DNA Repair Gene From Schizosaccharomyces pombe With Homology to Escherichia coli RecA. Genetics 152, 1557–1572.

Kick the bar chart habit, 2014. . Nat. Methods 11, 113–113. doi:10.1038/nmeth.2837

Kim, S., Schwille, P., 2003. Intracellular applications of fluorescence correlation spectroscopy: prospects for neuroscience. Curr. Opin. Neurobiol. 13, 583–590. doi:10.1016/j.conb.2003.09.002

Kohli, J., Hottinger, H., Munz, P., Strauss, A., Thuriaux, P., 1977. Genetic mapping in Schizosaccharomyces pombe by mitotic and meiotic analysis and induced haploidization. Genetics 87, 471–489.

Korlach, J., Schwille, P., Webb, W.W., Feigenson, G.W., 1999. Characterization of lipid bilayer phases by confocal microscopy and fluorescence correlation spectroscopy. Proc. Natl. Acad. Sci. U. S. A. 96, 8461–8466.

Krings, G., Bastia, D., 2004. swi1- and swi3-dependent and independent replication fork arrest at the ribosomal DNA of Schizosaccharomyces pombe. Proc. Natl. Acad. Sci. U. S. A. 101, 14085–14090. doi:10.1073/pnas.0406037101

Kumagai, A., Dunphy, W.G., 1999. Binding of 14-3-3 proteins and nuclear export control the intracellular localization of the mitotic inducer Cdc25. Genes Dev. 13, 1067–1072.

Kumar, S., Huberman, J.A., 2009. Checkpoint-Dependent Regulation of Origin Firing and Replication Fork Movement in Response to DNA Damage in Fission Yeast. Mol. Cell. Biol. 29, 602–611. doi:10.1128/MCB.01319-08

Lakowicz, J.R., 2006. Principles of fluorescence spectroscopy, 3rd ed. ed. Springer, New York.

Lambert, S., Carr, A.M., 2013. Impediments to replication fork movement: stabilisation, reactivation and genome instability. Chromosoma 122, 33–45. doi:10.1007/s00412-013-0398-9

Lambert, S., Mizuno, K. 'ichi, Blaisonneau, J., Martineau, S., Chanet, R., Fréon, K., Murray, J.M., Carr, A.M., Baldacci, G., 2010. Homologous Recombination Restarts Blocked Replication Forks at the Expense of Genome Rearrangements by Template Exchange. Mol. Cell 39, 346–359. doi:10.1016/j.molcel.2010.07.015

Lambert, S., Watson, A., Sheedy, D.M., Martin, B., Carr, A.M., 2005. Gross Chromosomal Rearrangements and Elevated Recombination at an Inducible Site-Specific Replication Fork Barrier. Cell 121, 689–702. doi:10.1016/j.cell.2005.03.022

Lammers, M., Follmann, H., 1984. Deoxyribonucleotide biosynthesis in yeast (Saccharomyces cerevisiae). Eur. J. Biochem. 140, 281–287. doi:10.1111/j.1432-1033.1984.tb08099.x

Lando, D., Endesfelder, U., Berger, H., Subramanian, L., Dunne, P.D., McColl, J., Klenerman, D., Carr, A.M., Sauer, M., Allshire, R.C., Heilemann, M., Laue, E.D., 2012. Quantitative single-molecule microscopy reveals that CENP-ACnp1 deposition occurs during G2 in fission yeast. Open Biol. 2, 120078–120078. doi:10.1098/rsob.120078

Larsson, K.-M., Jordan, A., Eliasson, R., Reichard, P., Logan, D.T., Nordlund, P., 2004. Structural mechanism of allosteric substrate specificity regulation in a ribonucleotide reductase. Nat. Struct. Mol. Biol. 11, 1142–1149. doi:10.1038/nsmb838

Latt, S.A., Stetten, G., Juergens, L.A., Willard, H.F., Scher, C.D., 1975. Recent developments in the detection of deoxyribonucleic acid synthesis by 33258 Hoechst fluorescence. J. Histochem. Cytochem. 23, 493–505. doi:10.1177/23.7.1095650

Lee, S.-H., Shin, J.Y., Lee, A., Bustamante, C., 2012. Counting single photoactivatable fluorescent molecules by photoactivated localization microscopy (PALM). Proc. Natl. Acad. Sci. 109, 17436–17441.

Lee, Y.D., Wang, J., Stubbe, J., Elledge, S.J., 2008. Dif1 is a DNA Damage Regulated Facilitator of Nuclear Import for Ribonucleotide Reductase. Mol. Cell 32, 70–80. doi:10.1016/j.molcel.2008.08.018

Leidenheimer, N.J., Harris, R.A., 1991. A transient osmotic permeabilization method for the introduction of impermeant molecules into functional brain membrane vesicles. J. Neurosci. Methods 40, 233–241.

Lelimousin, M., Adam, V., Nienhaus, G.U., Bourgeois, D., Field, M.J., 2009. Photoconversion of the Fluorescent Protein EosFP: A Hybrid Potential Simulation Study Reveals Intersystem Crossings. J. Am. Chem. Soc. 131, 16814–16823. doi:10.1021/ja905380y

Li, L., Liu, H., Dong, P., Li, D., Legant, W.R., Grimm, J.B., Lavis, L.D., Betzig, E., Tjian, R., Liu, Z., 2016. Real-time imaging of Huntingtin aggregates diverting target search and gene transcription. eLife 5. doi:10.7554/eLife.17056

Lichtman, J.W., Conchello, J.-A., 2005. Fluorescence microscopy. Nat. Methods 2, 910–919. doi:10.1038/nmeth817

Lieber, M.R., 2010. The Mechanism of Double-Strand DNA Break Repair by the Nonhomologous DNA End-Joining Pathway. Annu. Rev. Biochem. 79, 181–211. doi:10.1146/annurev.biochem.052308.093131

Lillemeier, B.F., Mörtelmaier, M.A., Forstner, M.B., Huppa, J.B., Groves, J.T., Davis, M.M., 2010. TCR and Lat are expressed on separate protein islands on T cell membranes and concatenate during activation. Nat. Immunol. 11, 90–96. doi:10.1038/ni.1832

Lindsay, H.D., Griffiths, D.J.F., Edwards, R.J., Christensen, P.U., Murray, J.M., Osman, F., Walworth, N., Carr, A.M., 1998. S-phase-specific activation of Cds1 kinase defines a subpathway of the checkpoint response in Schizosaccharomyces pombe. Genes Dev. 12, 382–395.

Liu, C., Powell, K., Mundt, K., Wu, L., Carr, A.M., Caspari, T., 2003. Cop9/signalosome subunits and Pcu4 regulate ribonucleotide reductase by both checkpoint-dependent and -independent mechanisms. Genes Dev. 17, 1130–1140. doi:10.1101/gad.1090803

Liu, Y., Li, M., Lee, E.Y.P., Maizels, N., 1999. Localization and dynamic relocalization of mammalian Rad52 during the cell cycle and in response to DNA damage. Curr. Biol. 9, 975–978.

Liu, Z., Lavis, L.D., Betzig, E., 2015. Imaging Live-Cell Dynamics and Structure at the Single-Molecule Level. Mol. Cell 58, 644–659. doi:10.1016/j.molcel.2015.02.033

Lopes, M., Cotta-Ramusino, C., Pellicioli, A., Liberi, G., Plevani, P., Muzi-Falconi, M., Newlon, C.S., Foiani, M., 2001. The DNA replication checkpoint response stabilizes stalled replication forks. Nature 412, 557–561. doi:10.1038/35087613

Lopez-Girona, A., Furnari, B., Mondesert, O., Russell, P., 1999. Nuclear localization of Cdc25 is regulated by DNA damage and a 14-3-3 protein. Nature 397, 172–175.

Los, G.V., Encell, L.P., McDougall, M.G., Hartzell, D.D., Karassina, N., Zimprich, C., Wood, M.G., Learish, R., Ohana, R.F., Urh, M., Simpson, D., Mendez, J., Zimmerman, K., Otto, P., Vidugiris, G., Zhu, J., Darzins, A., Klaubert, D.H., Bulleit, R.F., Wood, K.V., 2008. HaloTag: A Novel Protein Labeling Technology for Cell Imaging and Protein Analysis. ACS Chem. Biol. 3, 373–382. doi:10.1021/cb800025k

Lu, B.Y., Ma, J., Eissenberg, J.C., 1998. Developmental regulation of heterochromatin-mediated gene silencing in Drosophila. Development 125, 2223–2234.

Lundin, C., Erixon, K., Arnaudeau, C., Schultz, N., Jenssen, D., Meuth, M., Helleday, T., 2002. Different Roles for Nonhomologous End Joining and Homologous Recombination following Replication Arrest in Mammalian Cells. Mol. Cell. Biol. 22, 5869–5878. doi:10.1128/MCB.22.16.5869-5878.2002

Lundin, C., North, M., Erixon, K., Walters, K., Jenssen, D., Goldman, A.S.H., Helleday, T., 2005. Methyl methanesulfonate (MMS) produces heat-labile DNA damage but no detectable in vivo DNA double-strand breaks. Nucleic Acids Res. 33, 3799–3811. doi:10.1093/nar/gki681

Lundin, C., Schultz, N., Arnaudeau, C., Mohindra, A., Hansen, L.T., Helleday, T., 2003. RAD51 is Involved in Repair of Damage Associated with DNA Replication in Mammalian Cells. J. Mol. Biol. 328, 521–535. doi:10.1016/S0022-2836(03)00313-9

Magde, D., Elson, E., Webb, W.W., 1972. Thermodynamic fluctuations in a reacting system—measurement by fluorescence correlation spectroscopy. Phys. Rev. Lett. 29, 705.

Magde, D., Elson, E.L., Webb, W.W., 1974. Fluorescence correlation spectroscopy. II. An experimental realization. Biopolymers 13, 29–61. doi:10.1002/bip.1974.360130103

Manley, S., Gillette, J.M., Patterson, G.H., Shroff, H., Hess, H.F., Betzig, E., Lippincott-Schwartz, J., 2008. High-density mapping of single-molecule trajectories with photoactivated localization microscopy. Nat. Methods 5, 155–157. doi:10.1038/nmeth.1176

Marguerat, S., Schmidt, A., Codlin, S., Chen, W., Aebersold, R., Bähler, J., 2012. Quantitative Analysis of Fission Yeast Transcriptomes and Proteomes in Proliferating and Quiescent Cells. Cell 151, 671–683. doi:10.1016/j.cell.2012.09.019

Martinho, R.G., Lindsay, H.D., Flaggs, G., DeMaggio, A.J., Hoekstra, M.F., Carr, A.M., Bentley, N.J., 1998. Analysis of Rad3 and Chk1 protein kinases defines different checkpoint responses. EMBO J. 17, 7239–7249. doi:10.1093/emboj/17.24.7239

Mathews, C.K., 2006. DNA precursor metabolism and genomic stability. FASEB J. 20, 1300–1314. doi:10.1096/fj.06-5730rev

McEvoy, A.L., Hoi, H., Bates, M., Platonova, E., Cranfill, P.J., Baird, M.A., Davidson, M.W., Ewers, H., Liphardt, J., Campbell, R.E., 2012. mMaple: A Photoconvertible Fluorescent Protein for Use in Multiple Imaging Modalities. PLoS ONE 7, e51314. doi:10.1371/journal.pone.0051314

McKinney, S.A., Murphy, C.S., Hazelwood, K.L., Davidson, M.W., Looger, L.L., 2009. A bright and photostable photoconvertible fluorescent protein. Nat. Methods 6, 131–133. doi:10.1038/nmeth.1296

Mehta, A., Haber, J.E., 2014. Sources of DNA Double-Strand Breaks and Models of Recombinational DNA Repair. Cold Spring Harb. Perspect. Biol. 6, a016428. doi:10.1101/cshperspect.a016428

Meister, P., Taddei, A., Vernis, L., Poidevin, M., Gasser, S.M., Baldacci, G., 2005. Temporal separation of replication and recombination requires the intra-S checkpoint. J. Cell Biol. 168, 537–544. doi:10.1083/jcb.200410006

Meseth, U., Wohland, T., Rigler, R., Vogel, H., 1999. Resolution of fluorescence correlation measurements. Biophys. J. 76, 1619–1631.

Meurisse, J., Bacquin, A., Richet, N., Charbonnier, J.-B., Ochsenbein, F., Peyroche, A., 2014. Hug1 is an intrinsically disordered protein that inhibits ribonucleotide reductase activity by directly binding Rnr2 subunit. Nucleic Acids Res. 42, 13174–13185. doi:10.1093/nar/gku1095

Michel, B., Ehrlich, S.D., Uzest, M., 1997. DNA double-strand breaks caused by replication arrest. EMBO J. 16, 430–438.

Mizuno, H., Mal, T.K., Tong, K.I., Ando, R., Furuta, T., Ikura, M., Miyawaki, A., 2003. Photo-induced peptide cleavage in the green-to-red conversion of a fluorescent protein. Mol. Cell 12, 1051–1058.

Mortensen, K.I., Churchman, L.S., Spudich, J.A., Flyvbjerg, H., 2010. Optimized localization analysis for single-molecule tracking and super-resolution microscopy. Nat. Methods 7, 377–381. doi:10.1038/nmeth.1447

Müller, J.D., Chen, Y., Gratton, E., 2000. Resolving heterogeneity on the single molecular level with the photon-counting histogram. Biophys. J. 78, 474–486.

Muris, D.F., Vreeken, K., Carr, A.M., Broughton, B.C., Lehmann, A.R., Lohman, P.H., Pastink, A., 1993. Cloning the RAD51 homologue of Schizosaccharomyces pombe. Nucleic Acids Res. 21, 4586–4591.

Mütze, J., Ohrt, T., Schwille, P., 2011. Fluorescence correlation spectroscopy in vivo. Laser Photonics Rev. 5, 52–67. doi:10.1002/lpor.200910041

Nan, X., Collisson, E.A., Lewis, S., Huang, J., Tamguney, T.M., Liphardt, J.T., McCormick, F., Gray, J.W., Chu, S., 2013. Single-molecule superresolution imaging allows quantitative analysis of RAF multimer formation and signaling. Proc. Natl. Acad. Sci. 110, 18519–18524. doi:10.1073/pnas.1318188110

Nestoras, K., Mohammed, A.H., Schreurs, A.S., Fleck, O., Watson, A.T., Poitelea, M., O'Shea, C., Chahwan, C., Holmberg, C., Kragelund, B.B., Nielsen, O., Osborne, M., Carr, A.M., Liu, C., 2010. Regulation of ribonucleotide reductase by Spd1 involves multiple mechanisms. Genes Dev. 24, 1145–1159. doi:10.1101/gad.561910

Neumann, E., Rosenheck, K., 1972. Permeability changes induced by electric impulses in vesicular membranes. J. Membr. Biol. 10, 279–290. doi:10.1007/BF01867861

New, J.H., Sugiyama, T., Zaitseva, E., Kowalczykowski, S.C., 1998. Rad52 protein stimulates DNA strand exchange by Rad51 and replication protein A. Nature 391, 407–10.

Norbury, C., Blow, J., Nurse, P., 1991. Regulatory phosphorylation of the p34cdc2 protein kinase in vertebrates. EMBO J. 10, 3321–3329.

Nordlund, P., Reichard, P., 2006. Ribonucleotide reductases. Annu Rev Biochem 75, 681–706.

O'Connell, M.J., Raleigh, J.M., Verkade, H.M., Nurse, P., 1997. Chk1 is a wee1 kinase in the G2 DNA damage checkpoint inhibiting cdc2 by Y15 phosphorylation. EMBO J. 16, 545–554. doi:10.1093/emboj/16.3.545

Ormö, M., Cubitt, A.B., Kallio, K., Gross, L.A., Tsien, R.Y., Remington, S.J., 1996. Crystal structure of the Aequorea victoria green fluorescent protein. Science 273, 1392–1395.

Palayret, M., Armes, H., Basu, S., Watson, A.T., Herbert, A., Lando, D., Etheridge, T.J., Endesfelder, U., Heilemann, M., Laue, E., Carr, A.M., Klenerman, D., Lee, S.F., 2015. Virtual-'Light-Sheet' Single-Molecule Localisation Microscopy Enables Quantitative Optical Sectioning for Super-Resolution Imaging. PLoS ONE 10, e0125438. doi:10.1371/journal.pone.0125438

Patel, R.C., Kumar, U., Lamb, D.C., Eid, J.S., Rocheville, M., Grant, M., Rani, A., Hazlett, T., Patel, S.C., Gratton, E., Patel, Y.C., 2002. Ligand binding to somatostatin receptors induces receptor-specific oligomer formation in live cells. Proc. Natl. Acad. Sci. U. S. A. 99, 3294–3299. doi:10.1073/pnas.042705099

Patterson, G.H., 2002. A Photoactivatable GFP for Selective Photolabeling of Proteins and Cells. Science 297, 1873–1877. doi:10.1126/science.1074952

Peng, J.C., Karpen, G.H., 2009. Heterochromatic Genome Stability Requires Regulators of Histone H3 K9 Methylation. PLoS Genet. 5, e1000435. doi:10.1371/journal.pgen.1000435

Petermann, E., Orta, M.L., Issaeva, N., Schultz, N., Helleday, T., 2010. Hydroxyurea-Stalled Replication Forks Become Progressively Inactivated and Require Two Different RAD51-Mediated Pathways for Restart and Repair. Mol. Cell 37, 492–502. doi:10.1016/j.molcel.2010.01.021

Planchon, T.A., Gao, L., Milkie, D.E., Davidson, M.W., Galbraith, J.A., Galbraith, C.G., Betzig, E., 2011. Rapid three-dimensional isotropic imaging of living cells using Bessel beam plane illumination. Nat. Methods 8, 417–423. doi:10.1038/nmeth.1586

Politz, J.C., Browne, E.S., Wolf, D.E., Pederson, T., 1998. Intranuclear diffusion and hybridization state of oligonucleotides measured by fluorescence correlation spectroscopy in living cells. Proc. Natl. Acad. Sci. 95, 6043–6048.

Prahl, S., 2014. Rhodamine B [WWW Document]. URL http://omlc.org/spectra/PhotochemCAD/html/009.html (accessed 6.17.16).

Puchner, E.M., Walter, J.M., Kasper, R., Huang, B., Lim, W.A., 2013. Counting molecules in single organelles with superresolution microscopy allows tracking of the endosome maturation trajectory. Proc. Natl. Acad. Sci. 110, 16015–16020. doi:10.1073/pnas.1309676110

Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B.D., Sun, Z.-W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D., Jenuwein, T., 2000. Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature 406, 593–599. doi:10.1038/35020506

Reichard, P., 1993. From RNA to DNA, why so many ribonucleotide reductases? Science 260, 1773–1777.

Reinert, K.C., Dunbar, R.L., Gao, W., Chen, G., Ebner, T.J., 2004. Flavoprotein Autofluorescence Imaging of Neuronal Activation in the Cerebellar Cortex In Vivo. J. Neurophysiol. 92, 199–211. doi:10.1152/jn.01275.2003

Renz, M., Daniels, B.R., Vamosi, G., Arias, I.M., Lippincott-Schwartz, J., 2012. Plasticity of the asialoglycoprotein receptor deciphered by ensemble FRET imaging and single-molecule counting PALM imaging. Proc. Natl. Acad. Sci. 109, E2989–E2997. doi:10.1073/pnas.1211753109

Ricci, M.A., Manzo, C., García-Parajo, M.F., Lakadamyali, M., Cosma, M.P., 2015. Chromatin Fibers Are Formed by Heterogeneous Groups of Nucleosomes In Vivo. Cell 160, 1145–1158. doi:10.1016/j.cell.2015.01.054

Ries, J., Bayer, M., Csúcs, G., Dirkx, R., Solimena, M., Ewers, H., Schwille, P., 2010. Automated suppression of sample-related artifacts in Fluorescence Correlation Spectroscopy. Opt. Express 18, 11073–11082.

Rigler, R., Mets, Ü., Widengren, J., Kask, P., 1993. Fluorescence correlation spectroscopy with high count rate and low background: analysis of translational diffusion. Eur. Biophys. J. 22, 169–175. doi:10.1007/BF00185777

Rigler, R., Pramanik, A., Jonasson, P., Kratz, G., Jansson, O.T., Nygren, P.-Å., Ståhl, S., Ekberg, K., Johansson, B.-L., Uhlén, S., Uhlén, M., Jörnvall, H., Wahren, J., 1999. Specific binding of proinsulin C-peptide to human cell membranes. Proc. Natl. Acad. Sci. U. S. A. 96, 13318–13323.

Rofougaran, R., Crona, M., Vodnala, M., Sjoberg, B.-M., Hofer, A., 2008. Oligomerization Status Directs Overall Activity Regulation of the Escherichia coli Class Ia Ribonucleotide Reductase. J. Biol. Chem. 283, 35310–35318. doi:10.1074/jbc.M806738200

Rofougaran, R., Vodnala, M., Hofer, A., 2006. Enzymatically Active Mammalian Ribonucleotide Reductase Exists Primarily as an 6beta2 Octamer. J. Biol. Chem. 281, 27705–27711. doi:10.1074/jbc.M605573200

Rottenfusser, R., Wilson, E.E., Davidson, M.W., 2016. ZEISS Microscopy Online Campus | Microscopy Basics | Fluorescence Microscopy [WWW Document]. URL http://zeiss-campus.magnet.fsu.edu/print/basics/fluorescence-print.html (accessed 7.7.16).

Rust, M.J., Bates, M., Zhuang, X., 2006a. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). Nat. Methods 3, 793–796. doi:10.1038/nmeth929

Rust, M.J., Bates, M., Zhuang, X., 2006b. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). Nat. Methods 3, 793–796. doi:10.1038/nmeth929

Rustici, G., Mata, J., Kivinen, K., Lió, P., Penkett, C.J., Burns, G., Hayles, J., Brazma, A., Nurse, P., Bähler, J., 2004. Periodic gene expression program of the fission yeast cell cycle. Nat. Genet. 36, 809–817. doi:10.1038/ng1377

Ryan, D.P., Owen-Hughes, T., 2011. Snf2-family proteins: chromatin remodellers for any occasion. Curr. Opin. Chem. Biol. 15, 649–656. doi:10.1016/j.cbpa.2011.07.022

Sabatinos, S.A., Green, M.D., Forsburg, S.L., 2012. Continued DNA Synthesis in Replication Checkpoint Mutants Leads to Fork Collapse. Mol. Cell. Biol. 32, 4986–4997. doi:10.1128/MCB.01060-12

Saintigny, Y., Delacôte, F., Varès, G., Petitot, F., Lambert, S., Averbeck, D., Lopez, B.S., 2001. Characterization of homologous recombination induced by replication inhibition in mammalian cells. EMBO J. 20, 3861–3870. doi:10.1093/emboj/20.14.3861

San Filippo, J., Sung, P., Klein, H., 2008. Mechanism of Eukaryotic Homologous Recombination. Annu. Rev. Biochem. 77, 229–257. doi:10.1146/annurev.biochem.77.061306.125255

Santos, A., Young, I.T., 2000. Model-Based Resolution: Applying the Theory in Quantitative Microscopy. Appl. Opt. 39, 2948. doi:10.1364/AO.39.002948

Sarabia, M.-J.F., McInerny, C., Harris, P., Gordon, C., Fantes, P., 1993. The cell cycle genes cdc22+ and suc22+ of the fission yeast Schizosaccharomyces pombe encode the large and small subunits of ribonucleotide reductase. Mol. Gen. Genet. MGG 238, 241–251.

Sauvageau, S., Stasiak, A.Z., Banville, I., Ploquin, M., Stasiak, A., Masson, J.-Y., 2005. Fission yeast rad51 and dmc1, two efficient DNA recombinases forming helical nucleoprotein filaments. Mol. Cell. Biol. 25, 4377–4387. doi:10.1128/MCB.25.11.4377-4387.2005

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D.J., Hartenstein, V., Eliceiri, K., Tomancak, P., Cardona, A., 2012. Fiji: an open-source platform for biological-image analysis. Nat. Methods 9, 676–682. doi:10.1038/nmeth.2019

Sengupta, P., Garai, K., Balaji, J., Periasamy, N., Maiti, S., 2003. Measuring size distribution in highly heterogeneous systems with fluorescence correlation spectroscopy. Biophys. J. 84, 1977–1984.

Sengupta, P., Jovanovic-Talisman, T., Skoko, D., Renz, M., Veatch, S.L., Lippincott-Schwartz, J., 2011. Probing protein heterogeneity in the plasma membrane using PALM and pair correlation analysis. Nat. Methods 8, 969–975. doi:10.1038/nmeth.1704

Shimomura, O., 2005. The discovery of aequorin and green fluorescent protein. J. Microsc. 217, 3–15.

Shimomura, O., Johnson, F.H., Saiga, Y., 1962. Extraction, Purification and Properties of Aequorin, a Bioluminescent Protein from the Luminous Hydromedusan, Aequorea. J. Cell. Comp. Physiol. 59, 223–239. doi:10.1002/jcp.1030590302

Shroff, H., Galbraith, C.G., Galbraith, J.A., Betzig, E., 2008. Live-cell photoactivated localization microscopy of nanoscale adhesion dynamics. Nat. Methods 5, 417–423. doi:10.1038/nmeth.1202

Single-photon avalanche diode, 2014. . Wikipedia Free Encycl.
Skala, M.C., Riching, K.M., Gendron-Fitzpatrick, A., Eickhoff, J., Eliceiri, K.W., White, J.G., Ramanujam, N., 2007. In vivo multiphoton microscopy of NADH and FAD redox states, fluorescence lifetimes, and cellular morphology in precancerous epithelia. Proc. Natl. Acad. Sci. 104, 19494–19499.
Skilling, J., Bryan, R.K., 1984. Maximum Entropy Image Reconstruction - General Algorithm. Mon. Not. R. Astron. Soc. 211, 111. doi:10.1093/mnras/211.1.111
Spitzer, M., Wildenhain, J., 2016. BoxPlotR [WWW Document]. BoxPlotR Web-Tool Gener. Box Plots. URL http://boxplot.tyerslab.com/ (accessed 6.15.16).
Stern, M., Jensen, R., Herskowitz, I., 1984. Five SWI genes are required for expression of the HO gene in yeast. J. Mol. Biol. 178, 853–868.
Stokes, G.G., 1852. On the Change of Refrangibility of Light. Philos. Trans. R. Soc. Lond. 142, 463–562. doi:10.1098/rstl.1852.0022
Sung, P., 1997. Yeast Rad55 and Rad57 proteins form a heterodimer that functions with replication protein A to promote DNA strand exchange by Rad51 recombinase. Genes Dev. 11, 1111–1121.
Sung, P., Robberson, D.L., 1995. DNA strand exchange mediated by a RAD51-ssDNA nucleoprotein filament with polarity opposite to that of RecA. Cell 82, 453–461.
Szymborska, A., de Marco, A., Daigle, N., Cordes, V.C., Briggs, J.A.G., Ellenberg, J., 2013. Nuclear Pore Scaffold Structure Analyzed by Super-Resolution Microscopy and Particle Averaging. Science 341, 655–658. doi:10.1126/science.1240672
Tarnowski, B.I., Spinale, F.G., Nicholson, J.H., 1991. DAPI as a useful stain for nuclear quantitation. Biotech. Histochem. Off. Publ. Biol. Stain Comm. 66, 297–302.
Thompson, L.H., Schild, D., 2001. Homologous recombinational repair of DNA ensures mammalian chromosome stability. Mutat. Res. 477, 131–153.
Thompson, R.E., Larson, D.R., Webb, W.W., 2002. Precise nanometer localization analysis for individual fluorescent probes. Biophys. J. 82, 2775–2783.
Tokunaga, M., Imamoto, N., Sakata-Sogawa, K., 2008. Highly inclined thin illumination enables clear single-molecule imaging in cells. Nat. Methods 5, 159–161.
Tsutsui, Y., Morishita, T., Iwasaki, H., Toh, H., Shinagawa, H., 2000. A recombination repair gene of Schizosaccharomyces pombe, rhp57, is a functional homolog of the Saccharomyces cerevisiae RAD57 gene and is phylogenetically related to the human XRCC3 gene. Genetics 154, 1451–1461.
Ulbrich, M.H., Isacoff, E.Y., 2007. Subunit counting in membrane-bound proteins. Nat. Methods 4, 319–321. doi:10.1038/nmeth1024
Veatch, S.L., Machta, B.B., Shelby, S.A., Chiang, E.N., Holowka, D.A., Baird, B.A., 2012. Correlation Functions Quantify Super-Resolution Images and Estimate Apparent Clustering Due to Over-Counting. PLoS ONE 7, e31457. doi:10.1371/journal.pone.0031457
Walworth, N.C., Bernards, R., 1996. rad-dependent response of the chk1-encoded protein kinase at the DNA damage checkpoint. Science 271, 353–356.
Watson, A.T., Garcia, V., Bone, N., Carr, A.M., Armstrong, J., 2008. Gene tagging and gene replacement using recombinase-mediated cassette exchange in Schizosaccharomyces pombe. Gene 407, 63–74. doi:10.1016/j.gene.2007.09.024
Watson, A.T., Werler, P., Carr, A.M., 2011. Regulation of gene expression at the fission yeast Schizosaccharomyces pombe urg1 locus. Gene 484, 75–85. doi:10.1016/j.gene.2011.05.028
Weiss, M., Elsner, M., Kartberg, F., Nilsson, T., 2004. Anomalous Subdiffusion Is a Measure for Cytoplasmic Crowding in Living Cells. Biophys. J. 87, 3518–3524. doi:10.1529/biophysj.104.044263

Wiedenmann, J., Ivanchenko, S., Oswald, F., Schmitt, F., Röcker, C., Salih, A., Spindler, K.D., Nienhaus, G.U., 2004. EosFP, a fluorescent marker protein with UV-inducible green-to-red fluorescence conversion. Proc. Natl. Acad. Sci. U. S. A. 101, 15905–15910.

Wiedenmann, J., Ivanchenko, S., Oswald, F., Schmitt, F., Röcker, C., Salih, A., Spindler, K.-D., Nienhaus, G.U., 2004. EosFP, a fluorescent marker protein with UV-inducible green-to-red fluorescence conversion. Proc. Natl. Acad. Sci. U. S. A. 101, 15905–15910.

Willis, N.A., Zhou, C., Elia, A.E.H., Murray, J.M., Carr, A.M., Elledge, S.J., Rhind, N., 2016. Identification of S-phase DNA damage-response targets in fission yeast reveals conservation of damage-response networks. Proc. Natl. Acad. Sci. U. S. A. 113, E3676-3685. doi:10.1073/pnas.1525620113

Wold, M.S., 1997. Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. Annu. Rev. Biochem. 66, 61–92. doi:10.1146/annurev.biochem.66.1.61

Wood, V., Gwilliam, R., Rajandream, M.-A., Lyne, M., Lyne, R., Stewart, A., Sgouros, J., Peat, N., Hayles, J., Baker, S., Basham, D., Bowman, S., Brooks, K., Brown, D., Brown, S., Chillingworth, T., Churcher, C., Collins, M., Connor, R., Cronin, A., Davis, P., Feltwell, T., Fraser, A., Gentles, S., Goble, A., Hamlin, N., Harris, D., Hidalgo, J., Hodgson, G., Holroyd, S., Hornsby, T., Howarth, S., Huckle, E.J., Hunt, S., Jagels, K., James, K., Jones, L., Jones, M., Leather, S., McDonald, S., McLean, J., Mooney, P., Moule, S., Mungall, K., Murphy, L., Niblett, D., Odell, C., Oliver, K., O'Neil, S., Pearson, D., Quail, M.A., Rabbinowitsch, E., Rutherford, K., Rutter, S., Saunders, D., Seeger, K., Sharp, S., Skelton, J., Simmonds, M., Squares, R., Squares, S., Stevens, K., Taylor, K., Taylor, R.G., Tivey, A., Walsh, S., Warren, T., Whitehead, S., Woodward, J., Volckaert, G., Aert, R., Robben, J., Grymonprez, B., Weltjens, I., Vanstreels, E., Rieger, M., Schäfer, M., Müller-Auer, S., Gabel, C., Fuchs, M., Düsterhöft, A., Fritzc, C., Holzer, E., Moestl, D., Hilbert, H., Borzym, K., Langer, I., Beck, A., Lehrach, H., Reinhardt, R., Pohl, T.M., Eger, P., Zimmermann, W., Wedler, H., Wambutt, R., Purnelle, B., Goffeau, A., Cadieu, E., Dréano, S., Gloux, S., Lelaure, V., Mottier, S., Galibert, F., Aves, S.J., Xiang, Z., Hunt, C., Moore, K., Hurst, S.M., Lucas, M., Rochet, M., Gaillardin, C., Tallada, V.A., Garzon, A., Thode, G., Daga, R.R., Cruzado, L., Jimenez, J., Sánchez, M., del Rey, F., Benito, J., Domínguez, A., Revuelta, J.L., Moreno, S., Armstrong, J., Forsburg, S.L., Cerutti, L., Lowe, T., McCombie, W.R., Paulsen, I., Potashkin, J., Shpakovski, G.V., Ussery, D., Barrell, B.G., Nurse, P., Cerrutti, L., 2002. The genome sequence of Schizosaccharomyces pombe. Nature 415, 871–880. doi:10.1038/nature724

Woollard, A., Basi, G., Nurse, P., 1996. A novel S phase inhibitor in fission yeast. EMBO J. 15, 4603.

Wu, C.-C., Li, T.-K., Farh, L., Lin, L.-Y., Lin, T.-S., Yu, Y.-J., Yen, T.-J., Chiang, C.-W., Chan, N.-L., 2011. Structural basis of type II topoisomerase inhibition by the anticancer drug etoposide. Science 333, 459–462. doi:10.1126/science.1204117

Wu, Y., Kanchanawong, P., Zaidel-Bar, R., 2015. Actin-Delimited Adhesion-Independent Clustering of E-Cadherin Forms the Nanoscale Building Blocks of Adherens Junctions. Dev. Cell 32, 139–154. doi:10.1016/j.devcel.2014.12.003

Xu, K., Zhong, G., Zhuang, X., 2013. Actin, spectrin and associated proteins form a periodic cytoskeletal structure in axons. Science 339. doi:10.1126/science.1232251

Zeng, Y., Piwnica-Worms, H., 1999. DNA Damage and Replication Checkpoints in Fission Yeast Require Nuclear Exclusion of the Cdc25 Phosphatase via 14-3-3 Binding. Mol. Cell. Biol. 19, 7410–7419.

Zhang, M., Chang, H., Zhang, Y., Yu, J., Wu, L., Ji, W., Chen, J., Liu, B., Lu, J., Liu, Y., 2012. Rational design of true monomeric and bright photoactivatable fluorescent proteins. Nat. Methods 9, 727–729.

Zhang, M., Chang, H., Zhang, Y., Yu, J., Wu, L., Ji, W., Chen, J., Liu, B., Lu, J., Liu, Y., Zhang, J., Xu, P., Xu, T., 2012. Rational design of true monomeric and bright photoactivatable fluorescent proteins. Nat. Methods 9, 727–729. doi:10.1038/nmeth.2021

Zhao, X., Muller, E.G., Rothstein, R., 1998. A suppressor of two essential checkpoint genes identifies a novel protein that negatively affects dNTP pools. Mol. Cell 2, 329–340.

Zhong, G., He, J., Zhou, R., Lorenzo, D., Babcock, H.P., Bennett, V., Zhuang, X., 2014. Developmental mechanism of the periodic membrane skeleton in axons. eLife 3. doi:10.7554/eLife.04581

Zhong, Z.-H., Pramanik, A., Ekberg, K., Jansson, O.T., Jörnvall, H., Wahren, J., Rigler, R., 2001. Insulin binding monitored by fluorescence correlation spectroscopy. Diabetologia 44, 1184–1188. doi:10.1007/s001250100608

Zhou, B.-B.S., Elledge, S.J., 2000. The DNA damage response: putting checkpoints in perspective. Nature 408, 433–439.

Zofall, M., Grewal, S.I.S., 2006. Swi6/HP1 Recruits a JmjC Domain Protein to Facilitate Transcription of Heterochromatic Repeats. Mol. Cell 22, 681–692. doi:10.1016/j.molcel.2006.05.010

Zou, L., Elledge, S.J., 2003. Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. Science 300, 1542–1548. doi:10.1126/science.1083430

Zünkler, B.J., Wos-Maganga, M., Panten, U., 2004. Fluorescence microscopy studies with a fluorescent glibenclamide derivative, a high-affinity blocker of pancreatic beta-cell ATP-sensitive K+ currents. Biochem. Pharmacol. 67, 1437–1444. doi:10.1016/j.bcp.2003.12.011

8 Appendix

8.1 Cell lines used

cell line	produced by	leu 1-32	ura4-D18	ade 6-704	Kan	Hygro	Nat	mating type
501	Prof Tony Carr	-	-	-				+
AW 501	Dr Adam Watson	-		-				+
AW 310	Dr Adam Watson	-	-					+
AW309	Dr Adam Watson	-	-					-
spd1::Hygro	Dr Kostas Nestoras					+		
AW 459	Dr Adam Watson	-						+
rad3::Kan	Dr Thomas Etheridge	-	-	-	+			+
cds1::Nat	Dr Thomas Etheridge	-	-	-			+	-
chk1:Kan	Dr Ellen Tsang	-	-	-	+			-
swi6:Ura4	Dr Edgar Hartsuiker							h90
clr4::Nat	Dr Jo Murray						+	+
rrp1::Kan	Dr Dorota Dziadkowiec				+			+
rrp2::Kan	Dr Dorota Dziadkowiec				+			+

Table 8-1 List of cell lines used

8.2 Primers used

8.2.1 Tagging rrp2

		primer name	purpose	sequence
A	**9**	rrp2CHK1_F	sequencing tagged rrp1	CAGACATGAAATGGAAAGGAGAATTGC
A	**10**	rrp2CHK2_F	sequencing tagged rrp2	AGAATAAGCGTTATGGCTCGGA
A	**11**	rrp2CHK3_F	sequencing tagged rrp2	ACAAGAGTATCGAAGATAACGCTCC
A	**12**	rrp2CHK4_F	sequencing tagged rrp2	TATGTAGAGAGTGCTTAACCCATGT
A	**13**	rrp2CHK1_R	sequencing tagged rrp2	CTCCTCGCAAACTGAGCAC
A	**14**	rrp2CHK2_R	sequencing tagged rrp2	AGCATCCACGAGCAGCT
A	**15**	rrp2CHK3_R	sequencing tagged rrp2	GGTTCGGGTAGCTTGGATATCAT

A	**16**	rrp2CHK4_R	sequencing tagged rrp2	CGTATTTAAAGAGGACGATGAATCATGC
A	**19**	rrp2baseCHK_F	checking rrp2 base strain	GGTTCTGCTATCAATACCTACGCA
A	**20**	rrp2baseCHK_R	checking rrp2 base strain	CCGGTACCTCTTGTATTATGGAACTTTC
A	**23**	rrp2Sph1_F	amplifying rrp2 with SphI site	AAAAGCATGCCAGCTGCTTTGATACTGCACG
A	**24**	rrp2Sal1_R	amplifying rrp2 with SalI site	AAAAGTCGACTTATCGTGATGACATTCCAAATAAAAATGACAACTCC
A	**27**	rrp2base_F	amplifying *ura4* with 80 bp homology	AATTTAGGGGGAAACAACTTTAGAAGAAAGTAGAACATATAAAAACATTAATTAAAATATTTGTTTATTAAAACAATTTACGGATCCCCGGGTTAATTAA
A	**28**	rrp2base_R	amplifying *ura4* with 80 bp homology	ATGTGCGACGTTTTTTATTTCTATTCSTGAAAACATATTATATAATACATATTGGTTTCTAGAATTCGAGCTCGTTTAAAC
C	**1**	rrp2Sph1_F	preparing rrp2 for c-term cutting	aaaaGCATGCCAGCTGCTTTGATACTGCACG
C	**2**	mEos3_Xma1R	preparing mEos for rrp2 c-term ligation	aaaaCCCGGGTTAACGACGAGCGTTGTCAGG
C	**3**	rrp2CtermSal1_R	preparing rrp2 for c-term cutting	aaaaGTCGACCCCGGGTAAGCTAGCTCGTGATGACATTCCAAATAAAAATGACAACTCC
C	**4**	mEos3_Nhe1F	preparing mEos for rrp2 c-term ligation	aaaaGCTAGCACTGGTTCTACAGGATCAACCGGCTCCACGGGGTCGATGTCCGCTATTAAGCCTGACATG
D	**1**	rrp2CtermSal1_R2	preparing rrp2 for c-term cutting	aaaaGTCGACACTCCCGGGTAAGCTAGCTCGTGATGACATTCCAAATAAAAATGACAACTCC
G	**1**	Rrp2mEosCHK_F	for sequencing	GAGAGAACTAATAGATAGCGCTTTAGGGG
H	**1**	Rrp2_mEos_chk	as above	GAGAACTAATAGATAGCGCTTTAGGG
M	**1**	rrp2_mEosCHK_F	either side of rrp2 mEos insert site	TCACGTCTGAATACAAAGGAGTTGTCATT
M	**2**	rrp2_mEosCHK_R	either side of rrp2 mEos insert site	CGCTATCTATTAGTTCTCTCTTTCTGTCTT

Table 8-2 List of primers used for tagging rrp2 C-terminally with mEos3.1 using the Cre-lox system

8.2.2 Tagging rrp1

		primer name	purpose	sequence
A	**1**	rrp1CHK1_F	sequencing tagged rrp1	TGGCCCGTCGTCTTCAA

A	**2**	rrp1CHK2_F	sequencing tagged rrp1	AGTCCGTTCGACACGATCG
A	**3**	rrp1CHK3_f	sequencing tagged rrp1	GCTAAGGATATTTCCAGACCGCTA
A	**4**	rrp1CHK4_F	sequencing tagged rrp1	CATTAGTTGATGCAACTATCCTTCAAGC
A	**5**	rrp1CHK1_R	sequencing tagged rrp1	ATTGGGCACTTTGCTATGATCGTAG
A	**6**	rrp1CHK2_R	sequencing tagged rrp1	GGAACTTTAAGAGGCTGTAAAACTCATCA
A	**7**	rrp1CHK3_R	sequencing tagged rrp1	GGACTGGGAGTATTATGCTGAAGTAAT
A	**8**	rrp1CHK4_R	sequencing tagged rrp1	GGGACCTAGCAAATTGTTCGTCTA
A	**21**	rrp1Sac1_F	amplifying rrp1 with SacI site	AAAAGAGCTCCCGCCATATTGCAT
A	**22**	rrp1Sal1_R	amplifying rrp1 with Sal1 site	AAAAGTCGACTCATGAATTAAGCCCAAATAGATATAGCAACTCTTCT
K	**3**	rrp1_UP_F2	rrp1 fusion pcr up fragment	CGTTAGCAGAATCATTCATTCAAATCTCTCT
E	**2**	rrp1_UP_R	rrp1 fusion pcr up fragment	CGGTGAGATGCAATATGGCGG
E	**3**	rrp1_M_UP	rrp1 fusion pcr marker fragment	CCGCCATATTGCATCTCACCGCGGATCCCCGGGTTAATTAA
K	**6**	Rrp1_DO_F2	rrp1 fusion pcr down fragment	TTATTTAATGTTAACGTTACTCAAATGAACTATACTCTACTG
E	**5**	rrp1_DO_R	rrp1 fusion pcr down fragment	GTGCATGCCATTTGGCTAAACAAGT
E	**6**	rrp1Sac1_F2	rrp1 fusion pcr	aaaaGAGCTCCTCATCTGCGCATTTCATCAATTGC
K	**5**	Rrp1_M_DO2	rrp1 fusion pcr down fragment	CAGTAGAGTATAGTTCATTTGAGTAACGTTAACATTAAATAAGAATTCGAGCTCGTTTAAAC
K	**8**	Rrp1_baseCHK_F2	new for checking rrp1 base strain	GCAGAATCATTCATTCAAATCTCTCTTTCAC
K	**9**	Rrp1_baseCHK_R2	new for checking rrp1 base strain	TGCAAAGAGGCCGCCG
I	**1**	Rrp1_tagfus_F	forward fusion PCR primer to add in N-term restriction site	ATGGATTCATTGTCTGCATATCCTCC
I	**2**	mEos3_Nhe1_F	adding Nhe1 site to mEos	aaaaGCTAGCATGTCCGCTATTAAGCCTGACA
I	**3**	Rrp1_tagfus_R	reverse fusion PCR primer to add in N-term restriction site	GGAGGATATGCAGACAATGAATCCATGCTAGCGATGATTAAGAGCTAGGCTATAAAAGAGACA
I	**4**	mEos3_Nhe1_R	adding Nhe1 site to mEos	aaaaGCTAGCGGAGCCGGTTGATCCTGTAGAACCAGTACGACGAGCGTTGTCAGG

Table 8-3 List of primers used for tagging rrp1 N-terminally with mEos3.1 using the Cre-lox system

8.2.3 Sequencing primers from Dr Watson

name	purpose	sequence
PAL_F	sequencing in pAW8	ATGTGCTGCAAGGCGAT
PAL_R	sequencing in pAW9	CTCGTATGTTGTGTGGAATTGT
URA4_F	primer inside *ura4*	GCCAAAAATTACACAAGATAGAATGGATGTTTG
URA4_R	primer inside *ura4*	GTACAAAGCCAATGAAAGATGTATGTAGATG
MX6_R	check gene replacement using Mx6 plasmid	CGTCAAGACTGTCAAGGAGGG
MX6_F	check gene replacement using Mx6 plasmid	TGCGTCAATCGTATGTGAATGCT

Table 8-4 Primers supplied by Dr Watson for sequencing

Primers for sequencing the cloning of the gene and tagged gene into pAW8. Primers inside the ura4 gene for checking and sequencing the base strain. Primers for checking the presence of marker genes from the Mx6 plasmid, used for gene replacement.

8.3 Photo-activatable fluorescent protein properties

PAFP	λabs$_1$	λem$_1$	λabs$_2$	λem$_2$	λact	QY$_1$	QY$_2$	source	PE	PE reported
Kaede	508	518	572	580	350-400	0.88	0.33	Ando *et al*., 2002		
PA-GFP	400	516	504	517	433	0.13	0.79	Patterson *et al*., 2002		
Eos	506	516	571	581	405	0.70	0.55	Weidenmann *et al*., 2004		
dEos	506	516	569	581	405	0.66	0.60	Weidenmann *et al*., 2004		
mEos	506	516	571	581	405	0.64	0.62	Weidenmann *et al*., 2004		
Dronpa	503	518	dark	dark	490 (to 2) 400 (to 1)	0.85	/	Ando *et al*., 2004		
mKikGR	505	515	580	591		0.69	0.63	Tsutsui *et al*., 2005		
Dendra2	490	507	553	573	488	0.50	0.55	Gurskaya *et al*., 2006	82%	Durisic *et al*., 2014
mEos2	506	519	573	584	405	0.84	0.66	McKinney *et al*., 2009	60%	Durisic *et al*., 2014
PA-mCherry	404	ND	564	595		ND	0.64	Subach *et al*., 2009		
mClavGR1	486	503	565	582		0.84	0.56	Hoi *et al*., 2010		
mClavGR2	488	504	566	583		0.77	0.53	Hoi *et al*., 2010	53%	Durisic *et al*., 2014
mEos3.1	505	513	570	580	405	0.83	0.62	Zhang *et al*., 2012		
mEos3.2	506	516	572	580	405	0.84	0.55	Zhang *et al*., 2012	40%	Durisic *et al*., 2014
mMaple	488	504	566	583		0.74	0.56	McEvoy *et al*., 2012	74%	Durisic *et al*., 2014
tdEos	506	516	569	581	405	0.66	0.60	McKinney *et al*., 2009		

Table 8-5 Table of PA-FPs and properties

Absorption and emission maxima for various PA-FPs, along with quantum yield (QY) and photo-activation efficiency (PE). The majority switch from a green form to a red form with the exception of PA-GFP which switches from a dim to a bright form both of which are green and Dronpa which is reversibly switched between a green and a dark state.

8.4 Sequences of mEos2 and mEos3.1

8.4.1 *S. pombe* optimised mEos2 DNA sequence

```
ATGTCcGCtATtAAgCCtGACATGAAgATtAAgCTtCGtATGGAgGGtAACGTcAACGGtCACCACTTCG
TcATtGACGGtGACGGtACcGGtAAgCCtTTCGAgGGtAAgCAATCcATGGACCTtGAgGTcAAgGAgGG
tGGtCCtCTtCCtTTCGCtTTCGACATtCTtACcACcGCtTTCCACTACGGtAACCGtGTcTTCGCtAAg
TACCCtGACAACATtCAAGACTACTTCAAgCAATCcTTCCCtAAgGGtTACTCcTGGGAgCGtTCcCTtA
CcTTCGAgGACGGtGGtATtTGCATTGCtCGtAACGACATtACcATGGAgGGtGACACcTTCTACAACAA
gGTcCGtTTCTACGGtACcAACTTCCCtGCtAACGGtCCtGTcATGCAAAAgAAgACcCTtAAgtggGAg
CCtTCcACcGAgAAgATGTACGTcCGtGACGGtGTcCTtACcGGtGACATTCATATGGCtCTtCTtCTtG
AgGGtAACGCtCACTACCGtTGCGACTTCCGtACcACcTACAAgGCtAAgGAgAAgGGtGTcAAgCTtCC
tGGtTATCACTTCGTtGACCACTGCATtGAgATtCTtTCcCACGACAAgGACTACAACAAgGTcAAgCTt
TACGAgCACGCtGTcGCtCACTCcGGtCTtCCtGACAACGCtCGtCGtTAA
```

8.4.2 mEos2 protein sequence

```
MSAIKPDMKIKLRMEGNVNGHHFVIDGDGTGKPFEGKQSMDLEVKEGGPLPFAFDILTTAFHYGNRVFAK
YPDNIQDYFKQSFPKGYSWERSLTFEDGGICiARNDITMEGDTFYNKVRFYGTNFPANGPVMQKKTLKWE
PSTEKMYVRDGVLTGDihMALLLEGNAHYRCDFRTTYKAKEKGVKLPGyHFVDHCIEILSHDKDYNKVKL
YEHAVAHSGLPDNARR
```

8.4.3 *S. pombe* optimised mEos3.1 DNA sequence

```
ATGTCcGCtATtAAgCCtGACATGAAgATtAAgCTtCGtATGGAgGGtAACGTcAACGGtCACCACTTCG
TcATtGACGGtGACGGtACcGGtAAgCCtTTCGAgGGtAAgCAATCcATGGACCTtGAgGTcAAgGAgGG
tGGtCCtCTtCCtTTCGCtTTCGACATtCTtACcACcGCtTTCCACTACGGtAACCGtGTcTTCGCtAAg
TACCCtGACAACATtCAAGACTACTTCAAgCAATCcTTCCCtAAgGGtTACTCcTGGGAgCGtTCcCTtA
CcTTCGAgGACGGtGGtATtTGCAATGCtCGtAACGACATtACcATGGAgGGtGACACcTTCTACAACAA
gGTcCGtTTCTACGGtACcAACTTCCCtGCtAACGGtCCtGTcATGCAAAAgAAgACcCTtAAgtggGAg
CCtTCcACcGAgAAgATGTACGTcCGtGACGGtGTcCTtACcGGtGACGTTGAAATGGCtCTtCTtCTtG
AgGGtAACGCtCACTACCGtTGCGACTTCCGtACcACcTACAAgGCtAAgGAgAAgGGtGTcAAgCTtCC
tGGtGCTCACTTCGTtGACCACTGCATtGAgATtCTtTCcCACGACAAgGACTACAACAAgGTcAAgCTt
TACGAgCACGCtGTcGCtCACTCcGGtCTtCCtGACAACGCtCGtCGtTAA
```

8.4.4 mEos3.1 protein sequence

```
MSAIKPDMKIKLRMEGNVNGHHFVIDGDGTGKPFEGKQSMDLEVKEGGPLPFAFDILTTAFHYGNRVFAK
YPDNIQDYFKQSFPKGYSWERSLTFEDGGICNARNDITMEGDTFYNKVRFYGTNFPANGPVMQKKTLKWE
PSTEKMYVRDGVLTGDVEMALLLEGNAHYRCDFRTTYKAKEKGVKLPGAHFVDHCIEILSHDKDYNKVKL
YEHAVAHSGLPDNARR
```

8.4.5 *S. pombe* optimised mEos3.2 DNA sequence

```
ATGTCcGCtATtAAgCCtGACATGAAgATtAAgCTtCGtATGGAgGGtAACGTcAACGGtCACCACTTCG
TcATtGACGGtGACGGtACcGGtAAgCCtTTCGAgGGtAAgCAATCcATGGACCTtGAgGTcAAgGAgGG
tGGtCCtCTtCCtTTCGCtTTCGACATtCTtACcACcGCtTTCCACTACGGtAACCGtGTcTTCGCtAAg
TACCCtGACAACATtCAAGACTACTTCAAgCAATCcTTCCCtAAgGGtTACTCcTGGGAgCGtTCcCTtA
CcTTCGAgGACGGtGGtATtTGCAATGCtCGtAACGACATtACcATGGAgGGtGACACcTTCTACAACAA
gGTcCGtTTCTACGGtACcAACTTCCCtGCtAACGGtCCtGTcATGCAAAAgAAgACcCTtAAgtggGAg
CCtTCcACcGAgAAgATGTACGTcCGtGACGGtGTcCTtACcGGtGACATTGAAATGGCtCTtCTtCTtG
```

AgGGtAACGCtCACTACCGtTGCGACTTCCGtACcACcTACAAgGCtAAgGAgAAgGGtGTcAAgCTtCC
tGGtGCTCACTTCGTtGACCACTGCATtGAgATtCTtTCcCACGACAAgGACTACAACAAgGTcAAgCTt
TACGAgCACGCtGTcGCtCACTCcGGtCTtCCtGACAACGCtCGtCGtTAA

8.4.6 mEos3.2 protein sequence

MSAIKPDMKIKLRMEGNVNGHHFVIDGDGTGKPFEGKQSMDLEVKEGGPLPFAFDILTTAFHYGNRVFAK
YPDNIQDYFKQSFPKGYSWERSLTFEDGGICNARNDITMEGDTFYNKVRFYGTNFPANGPVMQKKTLKWE
PSTEKMYVRDGVLTGDIEMALLLEGNAHYRCDFRTTYKAKEKGVKLPGAHFVDHCIEILSHDKDYNKVKL
YEHAVAHSGLPDNARR

8.5 Counting simulation results

8.5.1 Limited N cluster analysis

Precision	n	p	Density	Fitted N	Fitted P	Fit	Fraction
10	2	0.3	23.55	2	0.3015	*	16.57
10	2	0.4	60.88	2	0.4353	*	24.15
10	2	0.5	119	2	0.5412	*	30.13
10	2	0.6	349.5	2	0.652	*	45.01
10	2	0.7	407.4	2	0.7042	*	40.73
10	3	0.3	68.05	3	0.3238	*	18.25
10	3	0.4	113.7	3	0.4266	*	18.46
10	3	0.5	279.2	3	0.544	*	24.75
10	3	0.6	522.4	3	0.6379	*	29.31
10	3	0.7	610.4	3	0.6983	*	25.53
10	4	0.3	113.6	4	0.3247	*	16.37
10	4	0.4	191.1	4	0.4302	*	16.29
10	4	0.5	464.9	4	0.5459	*	22.22
10	4	0.6	696.6	4	0.6268	*	22.2
10	4	0.7	812.3	4	0.6954	*	19.41
20	2	0.3	4.68	2	0.3018	*	12.26
20	2	0.4	12.5	2	0.4257	*	18.8
20	2	0.5	38.96	2	0.5448	*	37.07
20	2	0.6	91.08	2	0.6526	*	44.24
20	2	0.7	405.2	2	0.7042	*	66.36
20	3	0.3	17.39	3	0.3298	*	16.7
20	3	0.4	32.34	3	0.4222	*	19.72
20	3	0.5	57.96	3	0.5351	*	22.36
20	3	0.6	213.6	3	0.6591	*	39.2
20	3	0.7	608.7	3	0.6835	*	53.58
20	4	0.3	32.43	4	0.3261	*	17.95
20	4	0.4	56.16	4	0.431	*	17.72

20	4	0.5	97.56	4	0.5358	*	20.29
20	4	0.6	446.3	4	0.6588	*	40.31
20	4	0.7	814.8	4	0.6753	*	47.81

8.5.2 Variable N cluster analysis

Precision	n	p	Density	Fitted N	Fitted P	Fit	Fraction
10	2	0.3	16.56	2	0.3206	*	10.78
10	2	0.4	31.17	2	0.4214	*	12.29
10	2	0.5	49.22	2	0.5236	*	13.68
10	2	0.6	91.56	2	0.6201	*	19.47
10	2	0.7	133.6	2	0.6956	*	20.07
10	3	0.3	20.73	3	0.302	*	6.267
10	3	0.4	32.78	3	0.4122	*	6.435
10	3	0.5	58.66	3	0.5084	*	7.7
10	3	0.6	87.91	3	0.6017	*	6.851
10	3	0.7	160.1	3	0.6937	*	8.716
10	4	0.3	18.75	4	0.304	*	3.387
10	4	0.4	43.23	4	0.3999	*	3.982
10	4	0.5	47.49	4	0.5004	*	3.152
10	4	0.6	84.3	4	0.6004	*	3.769
10	4	0.7	171.4	4	0.6988	*	4.874
20	2	0.3	2.281	2	0.3107	*	2.273
20	2	0.4	6.199	2	0.4219	*	11.27
20	2	0.5	11.69	2	0.5135	*	16.59
20	2	0.6	22.96	2	0.6154	*	18.79
20	2	0.7	38.18	2	0.6988	*	22.96
20	3	0.3	13.82	4	0.2408	-	14.35
20	3	0.4	9.18	3	0.4029	*	7.701
20	3	0.5	11.66	3	0.4995	*	5.435
20	3	0.6	21.34	3	0.6152	*	6.814
20	3	0.7	32.95	3	0.6984	*	8.009
20	4	0.3	14.29	5	0.2506	-	10.16
20	4	0.4	6.277	5	0.3216	-	2.319
20	4	0.5	15.89	4	0.5141	*	3.472
20	4	0.6	18.63	4	0.5973	*	2.824
20	4	0.7	32.69	4	0.6994	*	3.786

8.5.3 Simulating varying density

N	p	Density	Fitted_N	Fitted_P	Best	Fraction
2	0.3	2.328	2	0.2752	+	1.235
2	0.3	4.613	2	0.284	+	2.367
2	0.3	6.926	2	0.2815	+	5.747
2	0.3	9.441	2	0.3119	+	4.188
2	0.3	11.85	2	0.3234	+	6.68
2	0.3	14.24	2	0.332		10.22
2	0.3	16.56	2	0.3206	*	10.78
2	0.3	18.84	2	0.3318		12.51
2	0.3	20.98	3	0.1931		13.96
2	0.3	23.21	3	0.1733		16.57
2	0.3	29.24	3	0.1838		16.24
2	0.3	36.52	3	0.185		23.07
2	0.3	45.77	3	0.2095		23.84
2	0.3	57.03	3	0.2105		31.16
2	0.3	71.75	4	0.1557		34.31
2	0.3	89.51	4	0.1676		38.47
2	0.3	112.1	6	0.1146		44.16
2	0.3	139.7	8	0.09077		50.11
2	0.3	175.1	7	0.1142		55.9
2	0.4	3.055	2	0.3748	+	0.6849
2	0.4	6.242	2	0.4138	+	2.115
2	0.4	9.262	2	0.3964	+	2.929
2	0.4	12.56	2	0.411	+	5.06
2	0.4	15.62	2	0.4189	+	9.122
2	0.4	18.61	2	0.3977	+	8.591
2	0.4	21.86	2	0.4082	+	10.27
2	0.4	25.21	2	0.4195	+	11.28
2	0.4	27.98	3	0.2513		13.11
2	0.4	31.17	2	0.4214	*	12.29
2	0.4	38.78	3	0.2628		17.64
2	0.4	48.73	3	0.2642		20.39
2	0.4	61.08	3	0.2678		24.15
2	0.4	75.84	3	0.2756		27.26
2	0.4	95.49	3	0.2922		32.17
2	0.4	119.1	4	0.2227		36.84
2	0.4	148.6	5	0.185		41.87
2	0.4	186.4	7	0.1371		47.9
2	0.4	232.7	8	0.1298		52.73
2	0.5	3.871	2	0.4985	+	0
2	0.5	7.855	2	0.505	+	2.534
2	0.5	11.77	2	0.5032	+	4.315

2	0.5	15.8	2	0.4916	+	5.792
2	0.5	19.94	2	0.5077	+	6.573
2	0.5	23.64	2	0.5074	+	7.311
2	0.5	27.58	2	0.5179	+	8.247
2	0.5	31.04	2	0.5134	+	10.84
2	0.5	35.14	2	0.506	+	11.61
2	0.5	39.45	2	0.5118	+	12.68
2	0.5	49.22	2	0.5236	*	13.68
2	0.5	60.55	3	0.3392		18.68
2	0.5	76.23	3	0.3427		20.67
2	0.5	94.96	3	0.3525		25.25
2	0.5	119.6	3	0.3588		30.13
2	0.5	149.5	3	0.3762		34.62
2	0.5	186.7	4	0.2874		38.25
2	0.5	232.1	5	0.2408		43.34
2	0.5	290.5	7	0.1828		48.77
2	0.6	4.793	2	0.6079	+	1.072
2	0.6	9.121	2	0.5775	+	3.184
2	0.6	13.93	2	0.5948	+	3.37
2	0.6	18.83	2	0.6032	+	5.281
2	0.6	23.84	2	0.6007	+	5.922
2	0.6	28.4	2	0.6099	+	7.627
2	0.6	32.43	2	0.6021	+	9.23
2	0.6	37.3	2	0.5988	+	9.283
2	0.6	42.44	2	0.6023	+	9.342
2	0.6	47.03	2	0.6075	+	10.91
2	0.6	58.51	2	0.5991	+	13.46
2	0.6	73.48	2	0.6119	+	15.57
2	0.6	91.56	2	0.6201	*	19.47
2	0.6	114.2	3	0.4253		22.99
2	0.6	142.7	3	0.4307		26.88
2	0.6	178.9	3	0.4422		30.78
2	0.6	223.3	3	0.4554		35.11
2	0.6	280.5	4	0.3615		40.16
2	0.6	348.8	6	0.2528		45.01
2	0.7	5.531	2	0.6998	+	1.406
2	0.7	10.91	2	0.6942	+	2.134
2	0.7	16.34	2	0.6941	+	4.117
2	0.7	21.98	2	0.6953	+	4.29
2	0.7	27.23	2	0.6808	+	6.26
2	0.7	32.54	2	0.6912	+	6.289
2	0.7	38.58	2	0.6942	+	7.196
2	0.7	43.54	2	0.6904	+	8.237
2	0.7	49.38	2	0.6898	+	8.766
2	0.7	55.08	2	0.699	+	9.293

2	0.7	68.41	2	0.6939	+	11.46
2	0.7	85.26	2	0.6992	+	13.72
2	0.7	107.5	2	0.6989	+	17.01
2	0.7	133.6	2	0.6956	*	20.07
2	0.7	166.6	3	0.4979		24.38
2	0.7	208.8	3	0.5081		27.85
2	0.7	260.3	3	0.5155		31.47
2	0.7	326.1	4	0.4101		36.21
2	0.7	407	5	0.346		40.73
3	0.3	3.707	3	0.3374		0.722
3	0.3	6.996	3	0.2994	+	3.441
3	0.3	10.35	3	0.2888	+	2.833
3	0.3	13.98	3	0.2959	+	5.062
3	0.3	17.14	3	0.3024	+	6.378
3	0.3	20.73	3	0.302	*	6.267
3	0.3	24.42	4	0.2197		7.123
3	0.3	27.96	4	0.2272		8.7
3	0.3	31.62	4	0.2199		9.324
3	0.3	35.02	4	0.2287		9.427
3	0.3	43.88	4	0.2346		12.98
3	0.3	54.9	5	0.183		14.67
3	0.3	68.68	5	0.1903		18.25
3	0.3	84.84	5	0.1915		21.91
3	0.3	107.2	6	0.1644		25.05
3	0.3	133	8	0.1267		29.34
3	0.3	167.5	9	0.1201		34.07
3	0.3	211.1	10	0.1151		38.11
3	0.3	261.4	10	0.121		42.3
3	0.4	4.727	3	0.385	+	0
3	0.4	9.117	4	0.2842		1.741
3	0.4	14.2	3	0.4038	+	2.789
3	0.4	18.45	4	0.288		3.569
3	0.4	23.35	3	0.4021	+	4.439
3	0.4	28.49	4	0.3044		4.54
3	0.4	32.78	3	0.4122	*	6.435
3	0.4	37.48	4	0.3031		6.667
3	0.4	42.4	4	0.3036		6.99
3	0.4	46.82	4	0.3041		9.152
3	0.4	58.15	4	0.3056		10.5
3	0.4	73.27	4	0.3144		12.1
3	0.4	91.66	4	0.3196		14.96
3	0.4	114.3	4	0.3233		18.46
3	0.4	143.8	5	0.2633		21.46
3	0.4	179	6	0.2279		25.28
3	0.4	222.2	8	0.1755		29.43

3	0.4	278.9	11	0.1345		33.8
3	0.4	347.8	12	0.1305		38.44
3	0.5	5.73	3	0.4968	+	0.4967
3	0.5	11.49	3	0.4875	+	1.969
3	0.5	17.61	3	0.5039	+	2.049
3	0.5	23.65	3	0.4997	+	3.301
3	0.5	29.22	3	0.4963	+	3.409
3	0.5	34.99	4	0.3764		5.035
3	0.5	41.23	3	0.502	+	5.211
3	0.5	46.86	4	0.3822		5.581
3	0.5	52.95	3	0.5099	+	6.347
3	0.5	58.66	3	0.5084	*	7.7
3	0.5	73.89	4	0.3897		8.083
3	0.5	91.96	4	0.3902		10.72
3	0.5	114.2	4	0.3933		12.47
3	0.5	142.7	4	0.3992		16
3	0.5	179	5	0.3282		18.49
3	0.5	224.1	6	0.2813		21.84
3	0.5	279.1	6	0.2895		24.75
3	0.5	349.4	9	0.204		29.56
3	0.5	436.4	14	0.1384		33.5
3	0.6	7.145	3	0.6162	+	0.5605
3	0.6	13.92	3	0.5995	+	1.333
3	0.6	21.17	3	0.5931	+	1.761
3	0.6	28.01	3	0.5921	+	2.426
3	0.6	34.95	3	0.598	+	3.207
3	0.6	41.99	3	0.6031	+	3.288
3	0.6	49.46	3	0.6035	+	4.729
3	0.6	56.32	3	0.6014	+	4.48
3	0.6	62.93	3	0.5977	+	5.41
3	0.6	69.61	3	0.5982	+	6.015
3	0.6	87.91	3	0.6017	*	6.851
3	0.6	110.4	4	0.468		8.653
3	0.6	137.3	4	0.4687		10.96
3	0.6	171.6	4	0.4752		13.09
3	0.6	214.9	4	0.4826		15.35
3	0.6	268.1	5	0.3997		18.52
3	0.6	334.7	6	0.3417		21.86
3	0.6	419.5	7	0.307		25.51
3	0.6	523.9	13	0.1757		29.31
3	0.7	8.086	3	0.6881	+	0.364
3	0.7	16.58	3	0.6999	+	1.136
3	0.7	24.94	3	0.7028	+	1.644
3	0.7	33.19	3	0.704	+	2.273
3	0.7	40.72	3	0.6929	+	2.877

3	0.7	49.48	3	0.7002	+	3.385
3	0.7	57.22	3	0.692	+	3.415
3	0.7	65.55	3	0.6926	+	4.043
3	0.7	73.75	3	0.6963	+	4.558
3	0.7	82.17	3	0.6947	+	5
3	0.7	102.5	3	0.6986	+	6.056
3	0.7	127.7	3	0.6925	+	7.413
3	0.7	160.1	3	0.6937	*	8.716
3	0.7	199.9	4	0.5484		10.6
3	0.7	249.9	4	0.5534		13.13
3	0.7	313.5	4	0.5592		15.82
3	0.7	391.1	5	0.4647		18.34
3	0.7	488.8	6	0.4019		21.89
3	0.7	610.7	9	0.2856		25.53
4	0.3	4.641	3	0.4076		0.4651
4	0.3	9.469	4	0.3112	+	2.07
4	0.3	13.98	5	0.2349		2.189
4	0.3	18.75	4	0.304	*	3.387
4	0.3	23.26	6	0.1922		5.088
4	0.3	27.81	5	0.2332		5.148
4	0.3	32.82	5	0.2407		4.958
4	0.3	38.21	5	0.2505		6.237
4	0.3	42.85	5	0.2496		7.463
4	0.3	46.7	6	0.2038		8.391
4	0.3	57.75	5	0.2445		9.58
4	0.3	73.34	6	0.2059		10.66
4	0.3	91.77	7	0.1825		13.46
4	0.3	114	8	0.164		16.37
4	0.3	142.5	9	0.1466		20.25
4	0.3	178.7	11	0.1236		23.57
4	0.3	222.9	11	0.13		26.82
4	0.3	279.6	11	0.1362		31.05
4	0.3	349.3	11	0.146		35.14
4	0.4	6.207	5	0.3213		1.245
4	0.4	12.4	4	0.3993	+	1.74
4	0.4	18.78	5	0.3211		1.553
4	0.4	25	4	0.4041	+	2.801
4	0.4	30.83	5	0.315		2.673
4	0.4	37.66	5	0.324		4.359
4	0.4	43.23	4	0.3999	*	3.982
4	0.4	49.77	5	0.3225		5.3
4	0.4	56.05	5	0.3251		6.212
4	0.4	62.12	5	0.323		6.196
4	0.4	78.46	5	0.3301		7.165
4	0.4	97.04	5	0.3288		8.761

4	0.4	121.4	6	0.2787		11.5
4	0.4	153.1	7	0.2461		13.49
4	0.4	190.8	8	0.2208		16.29
4	0.4	237.2	9	0.2005		18.96
4	0.4	298.5	13	0.1454		22.39
4	0.4	371.8	17	0.1156		26.14
4	0.4	465.2	17	0.1219		30.24
4	0.5	7.898	4	0.5081	+	0.653
4	0.5	15.55	4	0.4994	+	0.962
4	0.5	23.13	4	0.4886	+	1.843
4	0.5	31.28	5	0.403		1.936
4	0.5	39.15	4	0.5044	+	2.178
4	0.5	47.49	4	0.5004	*	3.152
4	0.5	54.89	5	0.4083		3.158
4	0.5	62.42	5	0.4044		3.741
4	0.5	70.08	5	0.4086		4.904
4	0.5	78.34	5	0.4112		4.958
4	0.5	97.88	5	0.4113		6.709
4	0.5	121.8	5	0.4097		7.085
4	0.5	152.2	5	0.419		8.942
4	0.5	190.6	6	0.3571		10.73
4	0.5	237.5	7	0.309		13.57
4	0.5	298.9	8	0.2805		16.27
4	0.5	371.8	10	0.2312		18.79
4	0.5	465.3	16	0.1513		22.22
4	0.5	579.9	17	0.1487		25.66
4	0.6	9.34	4	0.5883	+	0.2166
4	0.6	18.95	4	0.6056	+	0.4961
4	0.6	28.14	4	0.6029	+	1.499
4	0.6	37.23	4	0.5894	+	1.744
4	0.6	46.71	5	0.4883		2.145
4	0.6	56.21	4	0.5978	+	2.272
4	0.6	65.33	4	0.596	+	3.003
4	0.6	74.68	4	0.5989	+	3.338
4	0.6	84.3	4	0.6004	*	3.769
4	0.6	93.32	5	0.4871		4.38
4	0.6	117.7	5	0.493		5.062
4	0.6	147.1	5	0.4979		6.234
4	0.6	183.3	5	0.4978		7.593
4	0.6	229	5	0.5022		9.272
4	0.6	284.8	6	0.4275		11.04
4	0.6	357	6	0.4326		13.04
4	0.6	446.1	8	0.3372		15.94
4	0.6	559.2	11	0.2573		19.04
4	0.6	697.4	19	0.157		22.2

4	0.7	10.85	4	0.689	+	0.2902
4	0.7	21.91	4	0.7017	+	0.8799
4	0.7	33.18	4	0.7051	+	0.9563
4	0.7	43.61	4	0.6952	+	1.358
4	0.7	54.59	4	0.6971	+	1.86
4	0.7	65.48	4	0.697	+	2.007
4	0.7	76.11	4	0.6961	+	2.158
4	0.7	86.76	4	0.6913	+	2.793
4	0.7	98.2	4	0.6969	+	2.951
4	0.7	109.2	4	0.6989	+	3.309
4	0.7	136	4	0.6954	+	4.331
4	0.7	171.4	4	0.6988	*	4.874
4	0.7	213.3	5	0.5745		6.523
4	0.7	267.2	5	0.5768		7.532
4	0.7	333.3	5	0.5813		9.12
4	0.7	416.8	5	0.5817		11.21
4	0.7	522.4	6	0.5025		13.46
4	0.7	652.4	9	0.3539		16.5
4	0.7	816.4	16	0.2111		19.41

8.6 Statistical power for Rrp2 motion blur experiments, chapter 5

	power (1-β)	sample size	mean (μ)	standard deviation (σ)	type I error criterion (α)
rad52 fig 5.9					
untreated	0.9115	126	234.12	216.4085	0.05
MMS	1.0000	162	324.48		
asynch separately fig 5.10a					
untreated	0.9989	70	155.59	99.027	0.05
MMS	0.9989	70	71.40		
asynch simultaneously fig 5.10b					
untreated	1.0000	124	228.40	149.15	0.05
MMS	1.0000	116	83.71		
untreated mon vs untreated bi fig 5.12a					
untreated	0.9521	75	129.39	71.977	0.05
MMS	0.9778	90	87.27		
untreated mon vs MMS mon fig 5.12d					
untreated	0.9999	75	129.39	68.856	0.05
MMS	1.0000	137	65.89		
untreated bi vs MMS bi fig5.12c					
untreated	0.9946	90	87.27	49.889	0.05
MMS	0.9972	99	53.76		
HU fig 5.13a					
untreated	0.9995	64	122.91	64.664	0.05
HU	0.9838	108	89.91		
IR fig 5.13b					
untreated	0.9457	46	239.72	141.47	0.05
IR	0.9826	60	133.93		
Δ*rad57* +/- MMS fig 5.14a					
untreated	0.8699	204	127.74	73.89954	0.05
MMS	0.9426	268	105.18		
Δ*rad57* fig 5.14b					
wt untreated	0.8013	95	157.12	68.84558	0.05
57 untreated	0.8553	110	128.53		
Δ*chk1* +/- MMS fig 5.15a					
untreated	0.9999	101	31.18	20.00351	0.05

MMS	1.0000	119	15.08		
Δ*cds1* +/- MMS fig 5.15c					
untreated	0.9883	59	102.47	45.17	0.05
MMS	0.9977	76	70.03		
Δ*rad3* +/- MMS fig 5.15e					
untreated	0.9998	98	69.68	41.06	0.05
MMS	1.0000	139	37.33		
Δ*clr4* +/- MMS fig 5.16a					
untreated	0.985	73	92.49	45.203	0.05
MMS	0.999	114	68.37		
Δ*swi6* +/- MMS fig 5.16b					
untreated	0.9493	36	92.08	59.25	0.05
MMS	0.8506	25	42.36		
Δ*rrp1* fig5.18a					
wt	0.9477	59	103.76	44.61	0.05
rrp1-	0.9998	142	74.55		

Table 8-6 Calculation of statistical power for Rrp2 motion blur experiments

With an α criterion of 0.05 a power level of > 0.8 is generally thought an acceptable balance between α (type I error, false positive) and β (type II error, false negative) risk.

8.7 Rrp1 N-terminal mEos3.1 sequencing data alignment

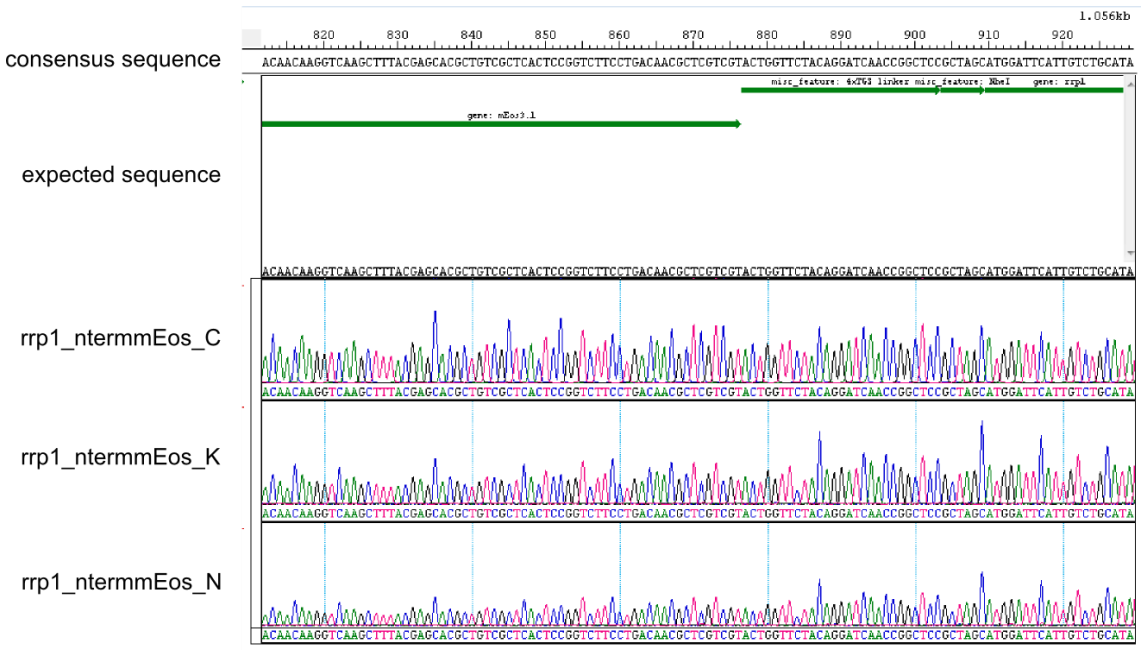


Figure 8.1 rrp1 mEos3.1 sequence alignment

Alignment showing the C-terminus of the mEos3.1 sequence and the N-terminus of the rrp1 sequence.